HUST

ĐẠI HỌC BÁCH KHOA HÀ NỘI

HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

ONE LOVE. ONE FUTURE.

ĐẠI HỌC
BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY
OF SCIENCE AND TECHNOLOGY

FORMATION TRACKING CONTROL FOR MULTI-ROBOT SYSTEMS BASED ON REINFORCEMENT LEARNING - ADP APPROACH

Instructor: Vũ Thị Thúy Nga

Student: Võ Khôi Thành Sơn – 20202768

ONE LOVE. ONE FUTURE.

Table of contents

HUST

hust.edu.vn fb.com/dhbkhn

Introduction

1. Introduction

- In this research, a system of multiple wheeled mobile robots is required to track and maintain specific formations while following desired trajectories.
- The objectives of this work are:
  - Introduce the multi-agent consensus control algorithms for single- and double-integrator models, and ensure formation tracking of wheeled mobile robot system in case of no disturbance by traditional control method.
  - Present an ADP-based adaptive optimal controller for the wheeled mobile robot formation tracking, considering wheel slip and external disturbances.

HUST

hust.edu.vn fb.com/dhbkhn

System modeling

A. Introduction of wheeled mobile robot (WMR)

- Two main driving wheels, one free wheel

Symbol	Meaning
$G(x_G, y_G)$	the center of mass of the platform
$M(x_M, y_M)$	the midpoint of the wheel shaft
F_1, F_2	total longitudinal friction forces at the right and left wheel
F_3	total lateral friction force along the wheel shaft
F_4, μ	external force and moment acting on G, respectively
γ_R, γ_L	longitudinal slip factors of the left and right wheels
η	the lateral slip factor along the wheel shaft
a	the distance between point M and G
b	half of the wheel shaft
θ	the orientation of the WMR
r	the radius of each wheel

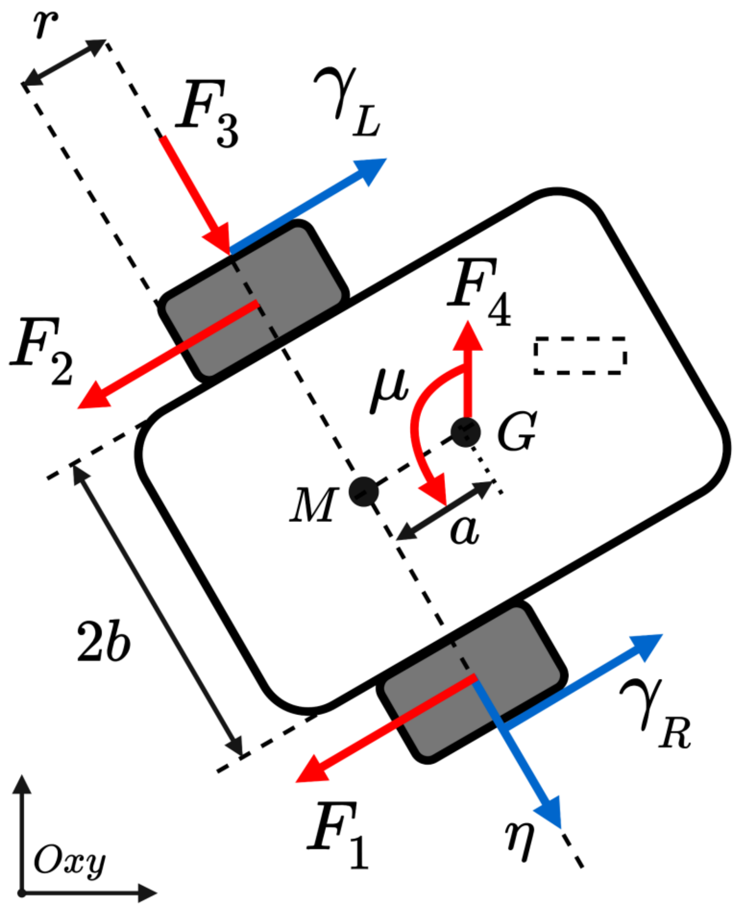


Diagram of the WMR

B. WMR system model:

1. Kinematic model

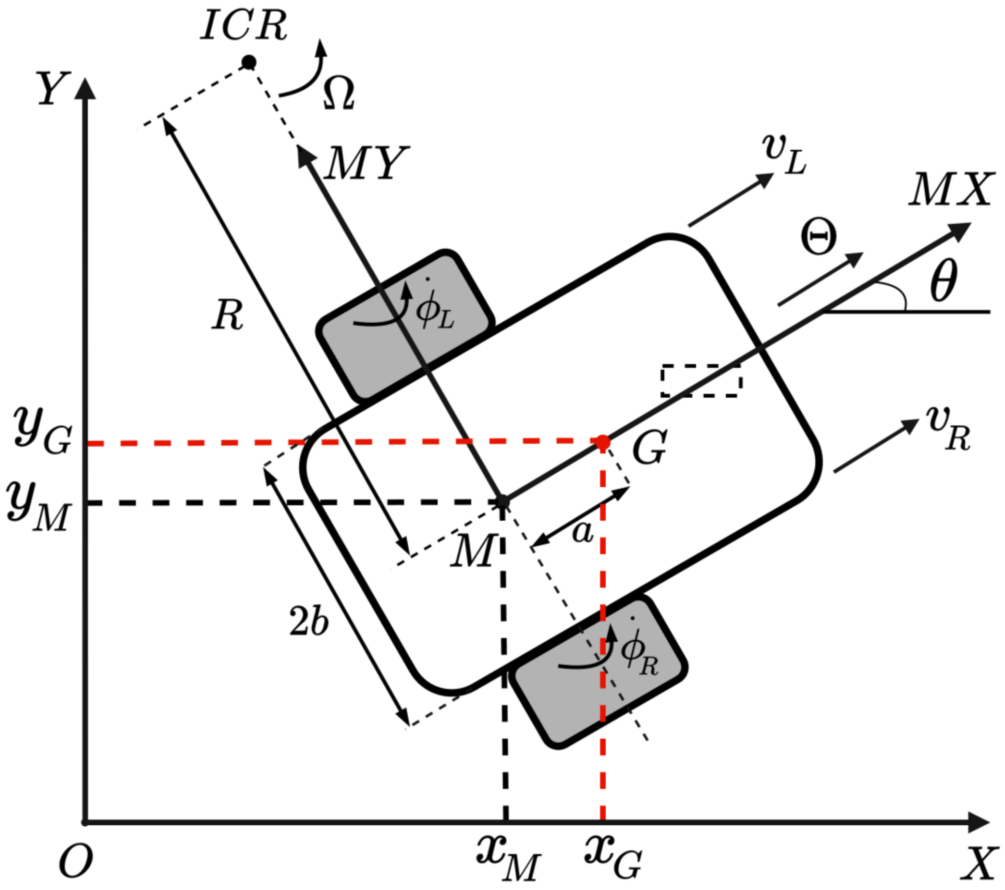
- In case of no disturbance:

The center of mass position is given by:

$$\begin{cases} x_G = x_M + a\cos(\theta) \\ y_G = y_M + a\sin(\theta) \end{cases} \quad (1)$$

The kinematics of the WMR is written as follows:

$$\begin{cases} \dot{x}_G = \Theta\cos(\theta) - \Omega a\sin(\theta) \\ \dot{y}_G = \Theta\sin(\theta) + \Omega a\cos(\theta) \quad (2) \\ \dot{\theta} = \Omega \end{cases}$$



WMR coordinates

Where Θ and Ω are the linear and angular velocities of the WMR:

$\Theta = \frac{r(\dot{\phi}_R+\dot{\phi}_L)}{2}$, $\Omega = \frac{r(\dot{\phi}_R-\dot{\phi}_L)}{2b}$, $\dot{\phi}_R$ and $\dot{\phi}_L$ are the angular velocities of the right and left wheel.

B. WMR system model:

1. Kinematic model

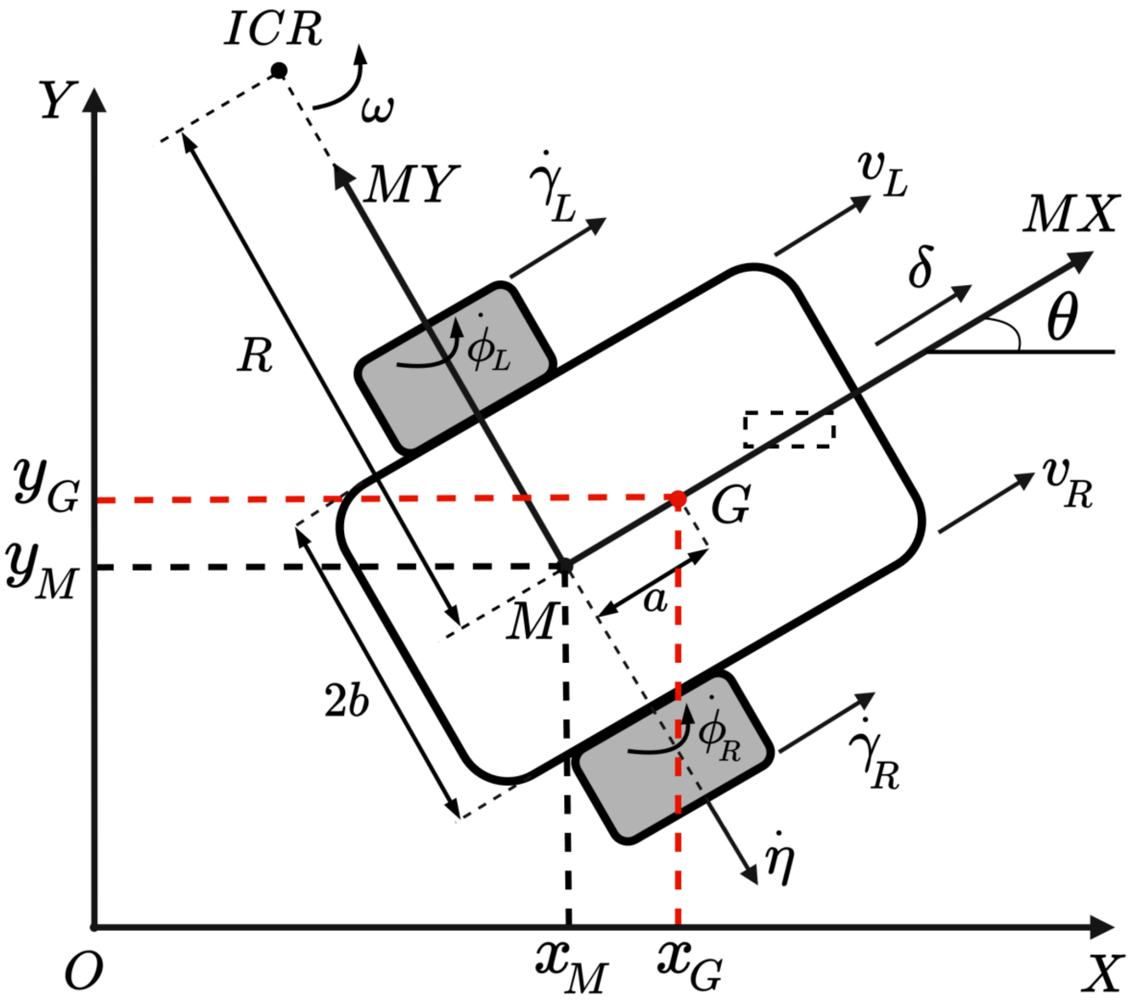
- In case of wheel slip and external disturbances

The actual linear and angular velocities are computed as:

$$\begin{cases} \delta = \dfrac{r(\dot{\phi}_R+\dot{\phi}_L)}{2} + \dfrac{\dot{\gamma}_R+\dot{\gamma}_L}{2} \\ \omega = \dfrac{r(\dot{\phi}_R-\dot{\phi}_L)}{2b} + \dfrac{\dot{\gamma}_R-\dot{\gamma}_L}{2b} \end{cases} \quad (3)$$

The actual kinematics of the WMR is written as follows:

$$\begin{cases} \dot{x}_G = \delta cos(\theta) - \omega a \sin(\theta) - \dot{\eta} \sin(\theta) \\ \dot{y}_G = \delta sin(\theta) + \omega a \cos(\theta) + \dot{\eta} \cos(\theta) \\ \dot{\theta} = \omega \end{cases} \quad (4)$$



WMR coordinates

B. WMR system model:

1. Kinematic model

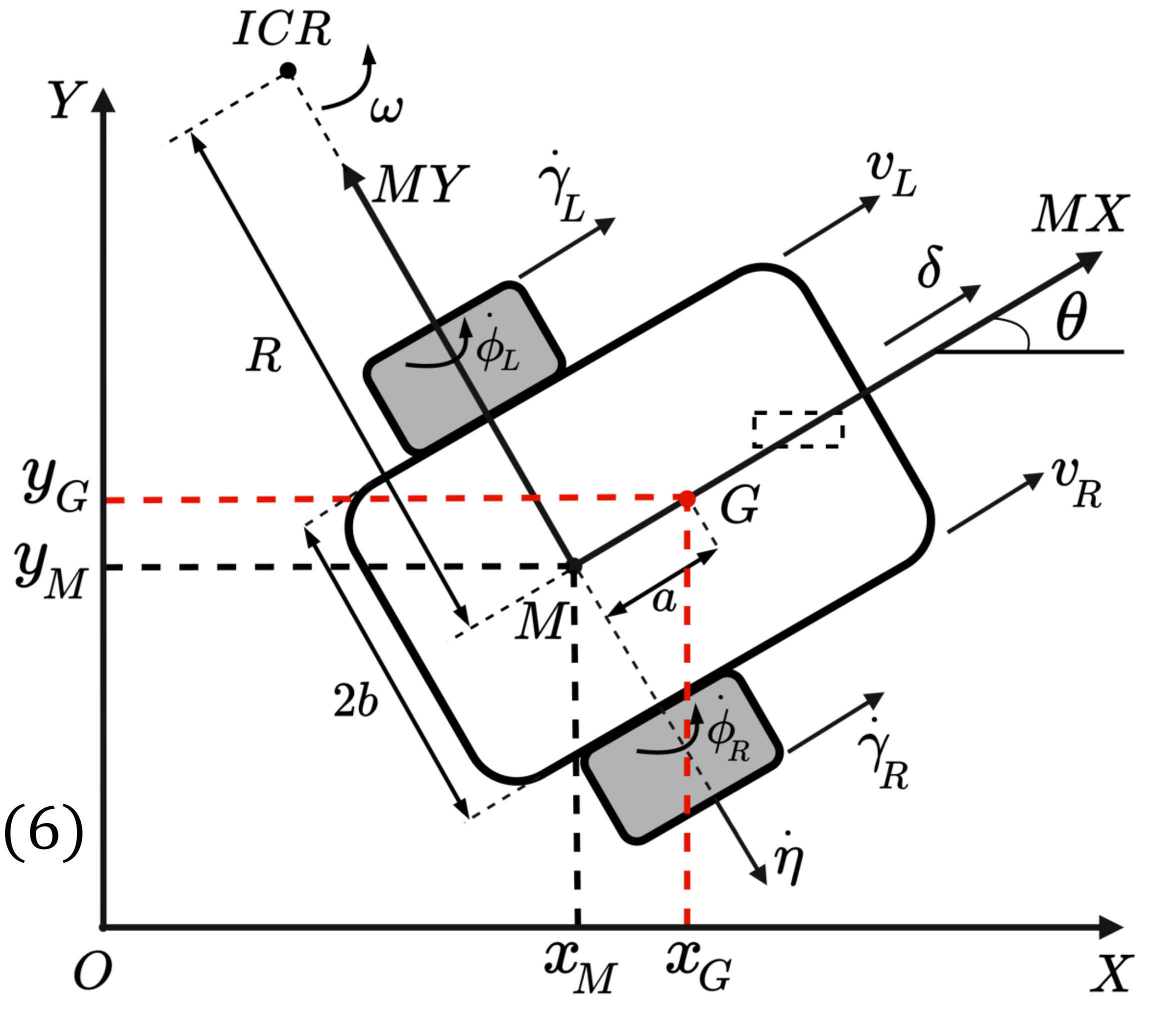
- In case of wheel slip and external disturbances

Combining (3) and (4), the position of the WMR can be rewritten as:

$$\dot{p} = Sv + d_p \qquad (5)$$

Where,

$$\begin{cases} p = [x_G \quad y_G]^T, \; v = [\dot{\phi}_R \quad \dot{\phi}_L]^T \\ S = \begin{bmatrix} \dfrac{r\cos(\theta)}{2} - \dfrac{ar\sin(\theta)}{2b} & \dfrac{r\cos(\theta)}{2} + \dfrac{ar\sin(\theta)}{2b} \\ \dfrac{r\sin(\theta)}{2} + \dfrac{ar\cos(\theta)}{2b} & \dfrac{r\sin(\theta)}{2} - \dfrac{ar\cos(\theta)}{2b} \end{bmatrix} \\ d_p = \begin{bmatrix} \dfrac{\cos(\theta)(\dot{\gamma}_R+\dot{\gamma}_L)}{2} - \dfrac{a\sin(\theta)(\dot{\gamma}_R-\dot{\gamma}_L)}{2b} - \dot{\eta}\sin(\theta) \\ \dfrac{\sin(\theta)(\dot{\gamma}_R+\dot{\gamma}_L)}{2} + \dfrac{a\cos(\theta)(\dot{\gamma}_R-\dot{\gamma}_L)}{2b} + \dot{\eta}\cos(\theta) \end{bmatrix} \end{cases} \qquad (6)$$



WMR coordinates

B. WMR system model:

2. Dynamic model

- In case of no disturbance

The dynamic model is derived by Lagrange formulation:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = F - A^T\lambda \quad (7)$$

Where,

+ L defines the Lagrangian (difference between kinetic and potential energy of the system), given as:

$$L = K - U \quad (8)$$

- $U = 0$: potential energy of the system
- K: kinetic energy of the WMR platform, derived as: $\mathrm{K} = K_G = \frac{1}{2}m_G(\dot{x}_G^2 + \dot{y}_G^2) + \frac{1}{2}I_G\Omega^2 \quad (9)$

m_G: mass of the platform of the WMR without the driving wheels

I_G: moment of inertia of the WMR platform about the vertical axis through G

B. WMR system model:

2. Dynamic model

- In case of no disturbance

The dynamic model is derived by Lagrange formulation:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = F - A^T\lambda \quad (7)$$

Where,

+ $\mathrm{q} = [x_G \quad y_G \quad \theta]^T$ is the general coordinate, $\lambda = [\lambda_1 \quad \lambda_2 \quad \lambda_3]^T$ is the Lagrange multiplier.

+ $A(q) = [-\sin(\theta) \quad \cos(\theta) \quad -a]$ is the matrix of kinematic constraint coefficients satisfying:

$$A(q)\dot{q} = 0 \quad (10)$$

+ $\mathrm{F} = [F_x \quad F_y \quad T]$ is the general force acting on the system

- $F_x = \frac{1}{r}(\tau_R + \tau_L)\cos(\theta)$: force in the x-axis direction
- $F_y = \frac{1}{r}(\tau_R + \tau_L)sin(\theta)$: force in the y-axis direction
- $T = \frac{b}{r}(\tau_R - \tau_L)$: the torque on the robot
- τ_R, τ_L: torques applied to the right and left wheels

B. WMR system model:

2. Dynamic model

- In case of no disturbance

According to Lagrange formulation (7), the following differential equations are obtained:

$$\begin{cases} m_G\ddot{x}_G = F_x + \lambda_1 \sin(\theta) \\ m_G\ddot{y}_G = F_y - \lambda_2 \cos(\theta) \\ I_G\ddot{\theta} = T + \lambda_3 a \end{cases} \qquad (11)$$

Or written in matrix form as: $$\bar{M}\ddot{q} = \bar{E}\tau - A^T\lambda \qquad (12)$$

Where,

$$\tau = [\tau_R \quad \tau_L]^T, \quad \bar{M} = \begin{bmatrix} m_G & 0 & 0 \\ 0 & m_G & 0 \\ 0 & 0 & I_G \end{bmatrix}, \quad \bar{E} = \frac{1}{\mathrm{r}}\begin{bmatrix} \cos(\theta) & \cos(\theta) \\ \sin(\theta) & \sin(\theta) \\ b & -b \end{bmatrix}$$

B. WMR system model:

2. Dynamic model

- In case of no disturbance

$$\bar{M}\ddot{q} = \bar{E}\tau - A^T\lambda \qquad (12)$$

The kinematics of the WMR (2) can be written in matrix form as: $\dot{q} = S_0 v_0 \qquad (13)$

$$\text{Where} \quad S_0 = \begin{bmatrix} \cos(\theta) & -a\sin(\theta) \\ \sin(\theta) & a\cos(\theta) \\ 0 & 1 \end{bmatrix}, \quad v_0 = [\Theta \quad \Omega]^T$$

Taking the time derivative of (13): $\ddot{q} = \dot{S}_0 v_0 + S_0 \dot{v}_0 \qquad (14)$

Substitute (14) into (12) : $\bar{M}\dot{S}_0 v_0 + \bar{M} S_0 \dot{v}_0 = \bar{E}\tau - A^T\lambda \qquad (15)$

Multiply (15) by S^T gives: $\boxed{M\dot{v}_0 + \bar{B} v_0 = E\tau} \qquad (16)$

$$\text{Where} \quad \mathrm{M} = \mathrm{S^T \bar{M} S} = \begin{bmatrix} m_G & 0 \\ 0 & m_G a^2 + I_G \end{bmatrix}, \quad \bar{B} = S^T \bar{M} \dot{S} = \begin{bmatrix} 0 & -m_G a\Omega \\ m_G a\Omega & 0 \end{bmatrix}, \quad E = S^T \bar{E} = \begin{bmatrix} \frac{1}{r} & \frac{1}{r} \\ \frac{b}{r} & -\frac{b}{r} \end{bmatrix}$$

B. WMR system model:

2. Dynamic model

- In case of wheel slip and external disturbances

The Lagrangian is defined as: $\mathrm{L} = K - U = K_G + K_R + K_L \quad (17)$

Where
$$\begin{cases} K_G = \frac{1}{2} m_G (\dot{x}_G^2 + \dot{y}_G^2) + \frac{1}{2} I_G \omega^2 & \text{: the kinetic energy of the WMR platform.} \\ K_R = \frac{1}{2} m_W \left[\left(r\dot{\phi}_R + \dot{\gamma}_R \right)^2 + \eta^2 \right] + \frac{1}{2} I_W \dot{\phi}_R^2 + \frac{1}{2} I_D \omega^2 & \text{: the kinetic energy of the right wheel.} \\ K_L = \frac{1}{2} m_W \left[\left(r\dot{\phi}_R + \dot{\gamma}_L \right)^2 + \eta^2 \right] + \frac{1}{2} I_W \dot{\phi}_L^2 + \frac{1}{2} I_D \omega^2 & \text{: the kinetic energy of the left wheel.} \end{cases}$$

- m_W: mass of each wheel
- I_W: the moment of inertia of each driving wheel about its rotational axis
- I_D: the moment of inertia of each driving wheel about its diameter axis

B. WMR system model:

2. Dynamic model

- In case of wheel slip and external disturbances

Based on (4) , the nonholonomic constraint equations can be written as:

$$\begin{cases} \dot{\gamma}_R = -r\dot{\phi}_R + \dot{x}_G\cos(\theta) + \dot{y}_G\sin(\theta) + b\omega \\ \dot{\gamma}_L = -r\dot{\phi}_L + \dot{x}_G\cos(\theta) + \dot{y}_G\sin(\theta) - b\omega \\ \dot{\eta} = -\dot{x}_G\sin(\theta) + \dot{y}_G\cos(\theta) \end{cases} \quad (18)$$

Similarly, the Lagrange formulation can be obtained as: $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} + \bar{\tau}_d = N\tau - A^T\lambda \quad (19)$

Where $\bar{\tau}_d$ illustrates model uncertainties and unknown disturbances, N is the input transformation matrix.

- $q = [x_G \quad y_G \quad \theta \quad \eta \quad \gamma_R \quad \gamma_L \quad \phi_R \quad \phi_L]^T$ is the general Lagrange coordinate
- $A(q) = \begin{bmatrix} \cos(\theta) & \sin(\theta) & b & 0 & -1 & 0 & -r & 0 \\ \cos(\theta) & \sin(\theta) & -b & 0 & 0 & -1 & 0 & -r \\ -\sin(\theta) & \cos(\theta) & a & -1 & 0 & 0 & 0 & 0 \end{bmatrix}$ which satisfies $A(q)\dot{q} = 0$

B. WMR system model:

2. Dynamic model

- In case of wheel slip and external disturbances

Solve the Lagrange equation gives: $\bar{M}\ddot{q} + \bar{\tau}_d = N\tau - A^T\lambda$ (20) where:

$$M = \begin{bmatrix} m_G & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & m_G & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_G + 2I_D & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2m_W & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & m_W & 0 & m_W r & 0 \\ 0 & 0 & 0 & 0 & 0 & m_W & 0 & m_W r \\ 0 & 0 & 0 & 0 & m_W r & 0 & m_W r^2 + I_W & 0 \\ 0 & 0 & 0 & 0 & 0 & m_W r & 0 & m_W r^2 + I_W \end{bmatrix}, \mathrm{N} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

B. WMR system model:

2. Dynamic model

- In case of wheel slip and external disturbances

The actual kinematics (6) can be written as: $\dot{q} = S_1(q)v + S_2(q)\dot{\gamma} + S_3(q)\,\dot{\eta}$ (21) where:

- $S_1 = \begin{bmatrix} \frac{r\cos(\theta)}{2} - \frac{ar\sin(\theta)}{2b} & \frac{r\cos(\theta)}{2} + \frac{ar\sin(\theta)}{2b} \\ \frac{r\sin(\theta)}{2} + \frac{ar\cos(\theta)}{2b} & \frac{r\sin(\theta)}{2} - \frac{ar\cos(\theta)}{2b} \\ \frac{r}{2b} & -\frac{r}{2b} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$, $S_2 = \begin{bmatrix} \frac{cos(\theta)}{2} - \frac{a\sin(\theta)}{2b} & \frac{cos(\theta)}{2} - \frac{a\sin(\theta)}{2b} \\ \frac{\sin(\theta)}{2} + \frac{a\cos(\theta)}{2b} & \frac{\sin(\theta)}{2} - \frac{a\cos(\theta)}{2b} \\ \frac{1}{2b} & -\frac{1}{2b} \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$, $S_3 = \begin{bmatrix} -\sin(\theta) \\ cos(\theta) \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$
- $v = [\dot{\phi}_R \quad \dot{\phi}_L]^T$, $\gamma = [\gamma_R \quad \gamma_L]^T$

B. WMR system model:

2. Dynamic model

- In case of wheel slip and external disturbances

$$\bar{M}\ddot{q} + \bar{\tau}_d = N\tau - A^T\lambda \qquad (20)$$

Taking the time derivative of (21): $\ddot{q} = \dot{S}_1 v + S_1\dot{v} + \dot{S}_2\dot{\gamma} + S_2\ddot{\gamma} + \dot{S}_3\dot{\eta} + S_3\ddot{\eta} \qquad (22)$

Substitute (22) into (20) and multiply both sides by S_1^T gives the actual dynamic model:

$$\boxed{M\dot{v} + Bv + Q\ddot{\gamma} + C\dot{\eta} + G\ddot{\eta} + \tau_d = \tau} \qquad (23)$$

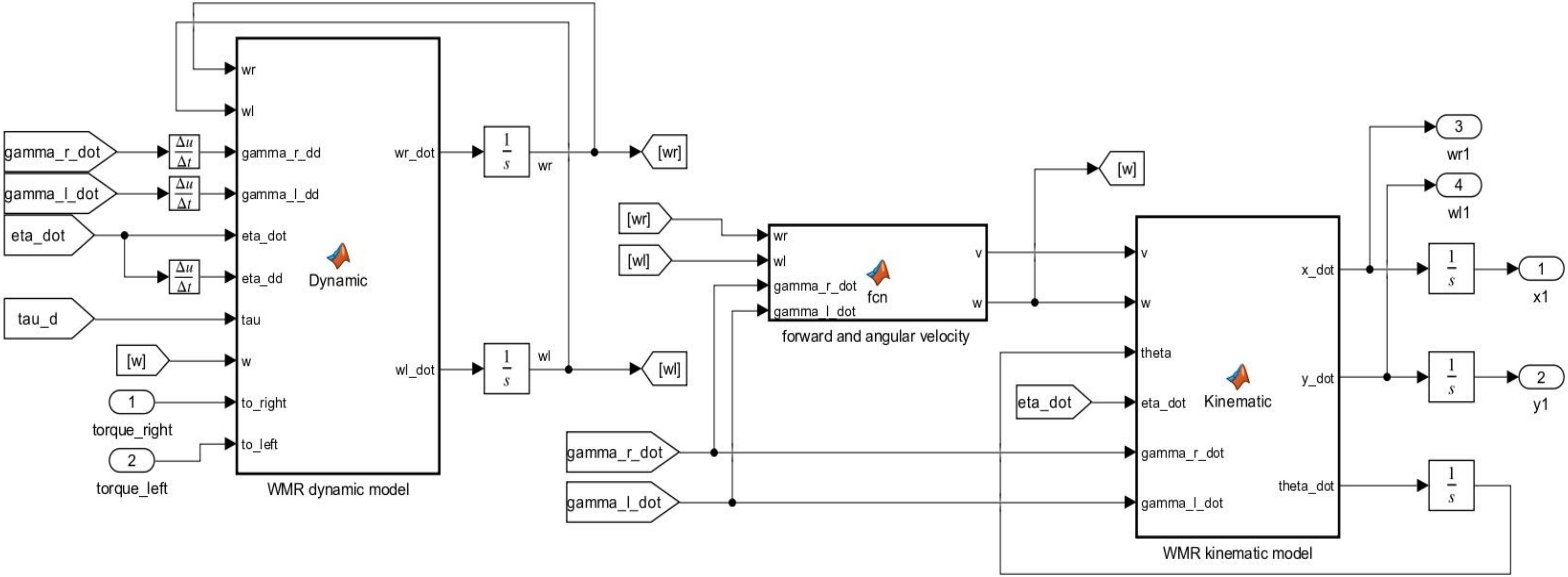
Where $\tau_d = S_1^T\bar{\tau}_d$; $M = S_1^T\bar{M}S_1 = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}$; $m_{22} = m_{11}, m_{21} = m_{12}$; $Q = S_1^T\bar{M}S_2 = \begin{bmatrix} Q_1 & Q_2 \\ Q_2 & Q_1 \end{bmatrix}$;

$$\mathrm{G} = \mathrm{S}_1^{\mathrm{T}}\bar{\mathrm{M}}\mathrm{S}_3 = \frac{m_G a r}{2b}\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad B = S_1^T\bar{M}\dot{S}_1 = \frac{m_G a r^2}{2b}\omega\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

$$m_{11} = m_G\left(\frac{r^2}{4} + \frac{a^2r^2}{4b^2}\right) + \frac{r^2}{4b^2}(I_G + 2I_D) + m_W r^2 + I_W \;, m_{12} = m_G\left(\frac{r^2}{4} - \frac{a^2r^2}{4b^2}\right) - \frac{r^2}{4b^2}(I_G + 2I_D)$$

$$Q_{1,2} = m_G\frac{r}{4}\left(1 \pm \frac{a^2}{b^2}\right) \pm \frac{r}{4b}(I_G + 2I_D)$$

C. WMR simulation implementation:



Multi-agent control design using traditional method

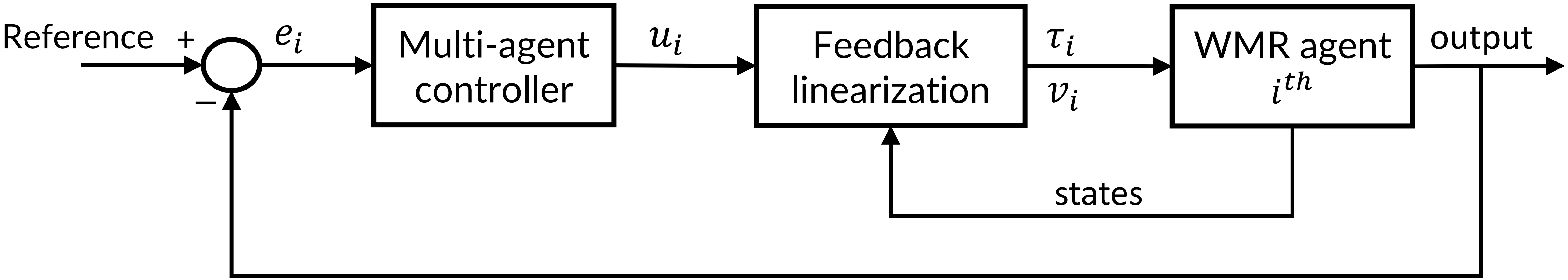
hust.edu.vn fb.com/dhbkhn

3. Multi-agent control using consensus algorithm

A. Control objective:

- The consensus algorithm is utilized for agents with single-integrator and double-integrator dynamics to achieve common information states.
- Distributed control law is implemented for each agent to ensure formation tracking while accurately following predefined trajectories.

➔ Apply the control algorithm to the WMR system.

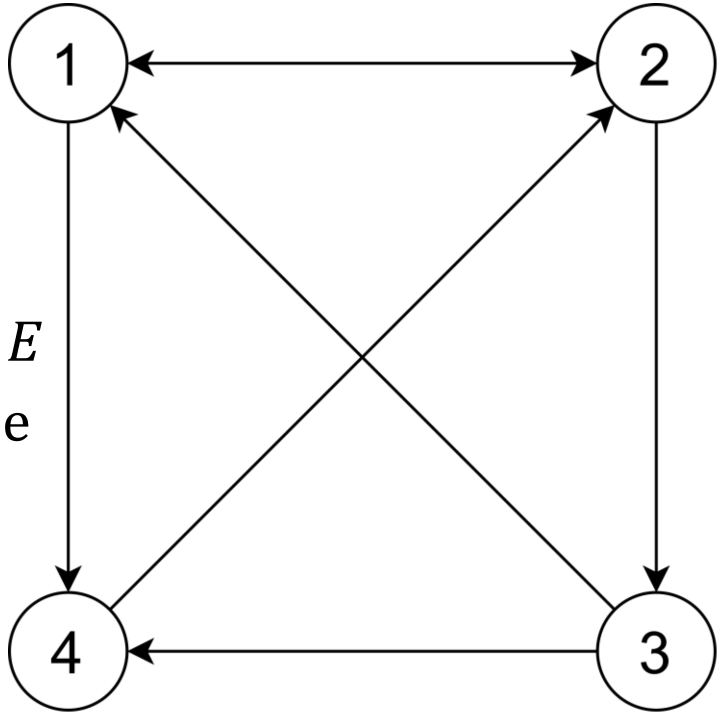


A. Control objective:

- Each agent interacts and exchanges relevant information with other agents through the communication graph.
- Consider a network of n agents with graph $G = \{V, E\}$ where:

- $V = \{1,2, \dots, n\}$ is a set of nodes
- $E \subset V \times V$ is a set of edges
- The adjacency matrix is $A = [a_{ij}]_{n\times n} \in \mathbb{R}^{n\times n}, \begin{cases} a_{ij} = 1 \; if \; (i,j) \in E \\ a_{ij} = 0, \text{otherwise} \end{cases}$
- The degree matrix is $\mathrm{D} = \mathrm{diag}\{\mathrm{d}_1, \dots, \mathrm{d}_\mathrm{n}\} \in \mathbb{R}^{n\times n}, \;\; d_i = \sum_{j=1}^{n} a_{ij}$
- The Laplacian matrix is $\mathcal{L} = D - A = [l_{ij}]_{n\times n} \in \mathbb{R}^{n\times n}$, where:

$$l_{ij} = \begin{cases} -a_{ij} = 1 \,, \; if \; i,j \in \mathrm{V}, \; \mathrm{i} \neq j \\ \sum_{j=1}^{n} a_{ij} \,, \; if \; i = j \end{cases}$$



Interaction graph of 4 agents

B. Consensus algorithm for single-integrator models:

Consider the case where each agent is modeled as single-integrator dynamics, expressed as:

$$\dot{p}_{ri} = u_i \qquad (24)$$

where p_{ri}, and u_i are the position and the control input of agent i^{th}, respectively.

The control law implemented for each agent:

$$u_i = \dot{p}_{ri}^d - k_p\left(p_{ri} - p_{ri}^d\right) + \sum_{j \in N_i} a_{ij}\left(\left(p_{rj} - p_{rj}^d\right) - \left(p_{ri} - p_{ri}^d\right)\right), i = 1, \dots, n \qquad (25)$$

- $\mathrm{N_i} = \{\mathrm{j} \in V | (i,j) \in E\}$: the set of neighborhood of agents i^{th}
- $p_{ri} = [x_{ri} \quad y_{ri}]^T$: actual position
- $p_{ri}^d = \left[x_{ri}^d \quad y_{ri}^d\right]^T$: desired position
- $u_i = [u_{xi} \quad u_{yi}]^T$: control input

Let $p_r = [p_{r1} \quad p_{r2} \quad \cdots \quad p_{rn}]^T$, $p_r^d = [p_{r1}^d \quad p_{r2}^d \quad \dots \quad p_{rn}^d]^T$, the control law is rewritten as:

$$u = -\left(\left(k_p I_n + \mathcal{L}\right) \otimes I_2\right)\left(p_r - p_r^d\right) \qquad (26)$$

Where $u = [u_1 \quad u_2 \quad \cdots \quad u_n]^T$, I_n is the identity matrix of size n

B. Consensus algorithm for single-integrator models:

The positions of each WMR agent under no disturbance (13) can be rewritten as:

$$\dot{p}_i = Sv_i \qquad (27)$$

Where $S = \begin{bmatrix} \cos(\theta) & -a\sin(\theta) \\ \sin(\theta) & a\cos(\theta) \end{bmatrix}, \quad v_i = [\Theta_i \quad \Omega_\mathrm{i}]^T$

Feedback linearization:

$$v_i = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\frac{1}{a}\sin(\theta) & \frac{1}{a}\cos(\theta) \end{bmatrix} u_i \qquad (28), \text{ where } u_i = [\mathrm{u}_{xi} \quad u_{y\mathrm{i}}]^T$$

The model becomes:

$$\dot{p}_i = u_i \qquad (29)$$

➔ Control law (26) can be directly applied.

C. Consensus algorithm for double-integrator models:

In case the dynamics of each agent follows double-integrator dynamics, defined as:

$$\begin{cases} \dot{p}_{ri} = v_{ri} \\ \dot{v}_{ri} = u_i \end{cases} \qquad (30)$$

where p_{ri} , v_{ri} and u_i are the position, velocity and the control input of agent i^{th}, respectively.

The control law implemented for each agent:

$$u_i = -k_p\left(p_{ri} - p_{ri}^d\right) + \sum_{j \in N_i} a_{ij}\left((p_{rj} - p_{rj}^d) - (p_{ri} - p_{ri}^d)\right) - k_v\left(v_{ri} - \dot{p}_{ri}^d\right), i = 1, \dots, n \quad (31)$$

Or rewritten as:

$$u = -\left((k_p I_n + \mathcal{L}) \otimes I_2\right)\left(p_r - p_r^d\right) - k_v\left(v_r - \dot{p}_r^d\right) \qquad (32)$$

C. Consensus algorithm for double-integrator models:

The dynamic model of each WMR agent under no disturbance (16) can be rewritten as:

$$\dot{v}_i = Bv_i + N\tau_i \quad (33)$$

$$\mathrm{B} = -\mathrm{M}^{-1}\overline{\mathrm{B}} = \begin{bmatrix} 0 & a\Omega_\mathrm{i} \\ \frac{-m_G a\Omega}{I_G + m_G a} & 0 \end{bmatrix}, \; N = M^{-1}E = \begin{bmatrix} \frac{1}{m_G r} & \frac{1}{m_G r} \\ \frac{b}{r(I_G + m_G a^2)} & \frac{-b}{r(I_G + m_G a^2)} \end{bmatrix}, \; v_i = \begin{bmatrix} \Theta_i \\ \Omega_i \end{bmatrix}, \; \tau_i = \begin{bmatrix} \tau_{Ri} \\ \tau_{Li} \end{bmatrix}$$

And the kinematic equations (27) obtained as: $\dot{p}_i = Sv_i$. Define: $z_i = Sv_i$ (34)

Derivative of z_i:

$$\dot{z}_i = \dot{S}v_i + S\dot{v}_i \quad (35)$$

Substitute (33) into (35):

$$\dot{z}_i = \left(\dot{S} + SB\right)v_i + SN\tau_i \quad (36)$$

Feedback linearization:

$$\tau_i = N^{-1}S^{-1}\left(u_i - \left(\dot{S} + SB\right)v_i\right) \quad (37)$$

The model becomes:

$$\begin{cases} \dot{p}_i = z_i \\ \dot{z}_i = u_i \end{cases} \quad (38)$$

➔ Control law (32) can be directly applied.

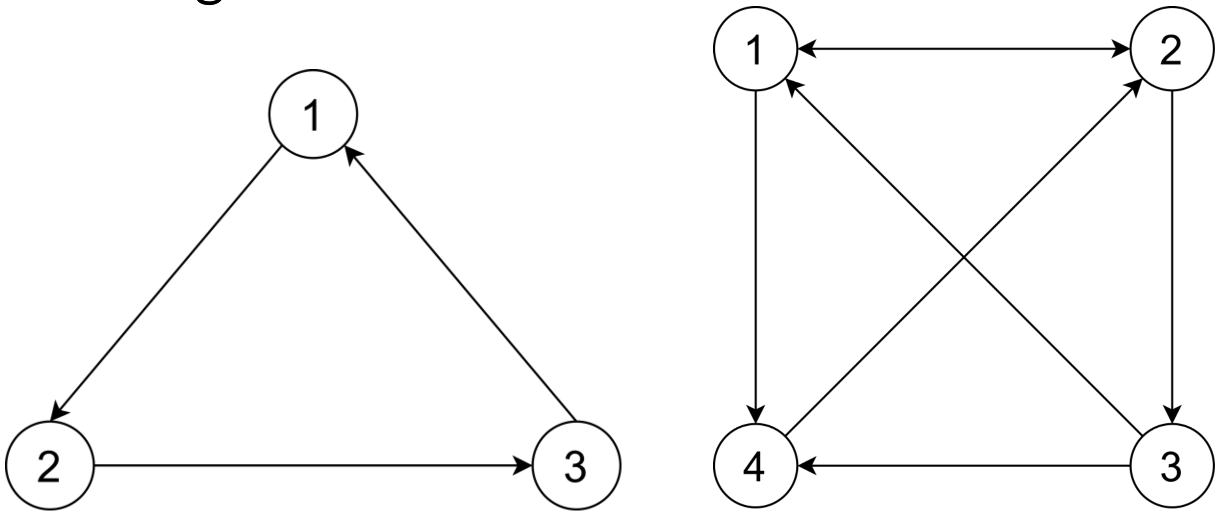
Multi-agent control design using Adaptive Dynamic Programming approach

hust.edu.vn fb.com/dhbkhn

4. Multi-agent control using ADP approach

A. Control objective:

- An adaptive optimal controller in cooperation with a disturbance observer is proposed to ensure formation tracking for WMR agents in case of wheel slip and external disturbances.
- A multi-robot system with n WMRs achieve formation geometry of square and triangle shape, following two types of trajectories: line and circle.
- For comparison with previous approach, the WMR formation is tasked to follow a curved trajectory under square configuration.



Interaction topology considered

4. Multi-agent control using ADP approach

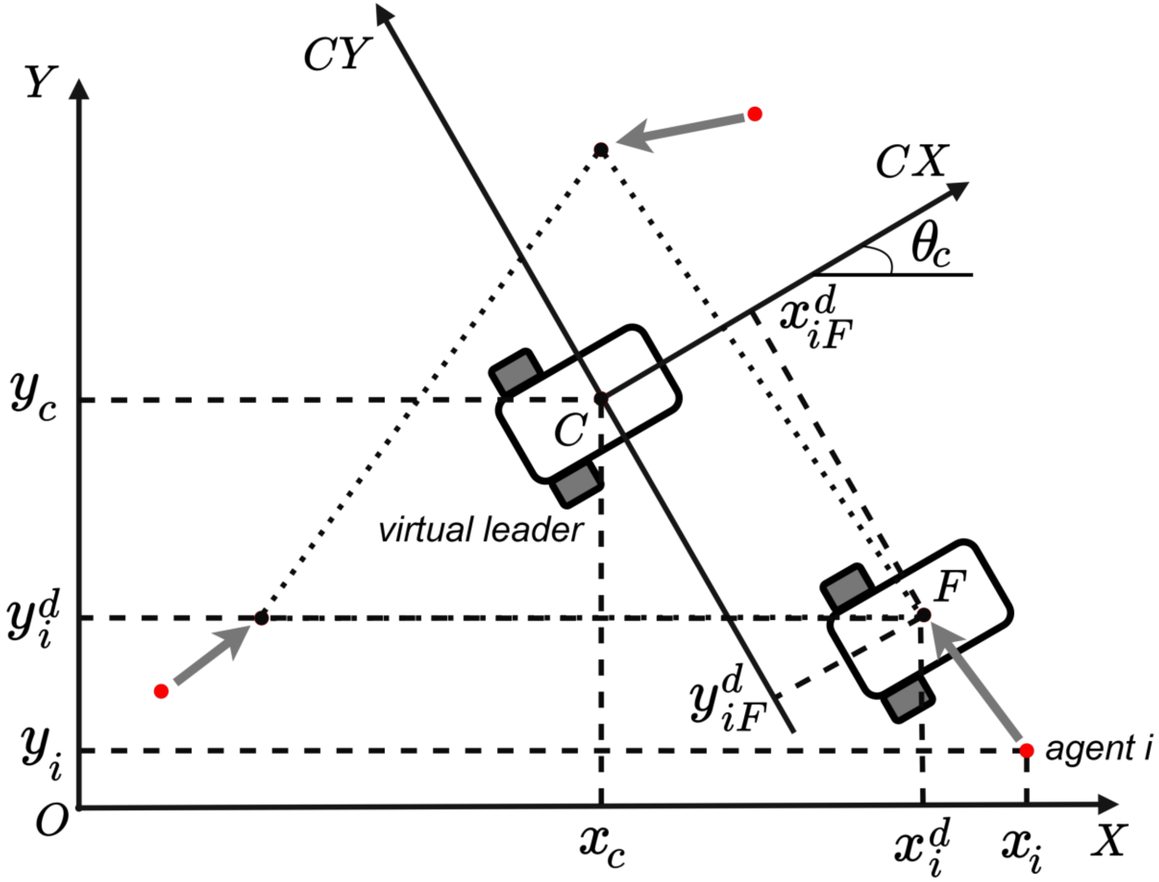
B. Problem formulation:

- Virtual leader/structure approach.
- Decentralized scheme instead of a centralized one.
- Control approach:

+ Define the desired dynamics of the virtual leader or virtual structure.

+ The motion of the virtual leader is translated into desired motions for each robot.

+ Derive distributed tracking control laws for each robot.



Triangle formation with a known virtual leader

4. Multi-agent control using ADP approach

B. Problem formulation:

The formation structure of the WMR system is derived as:

$$p_i^d = p_c + \begin{bmatrix} \cos(\theta_c) & -\sin(\theta_c) \\ \sin(\theta_c) & \cos(\theta_c) \end{bmatrix} p_{iF}^d \qquad (39)$$

- $p_i^d = \begin{bmatrix} x_i^d & y_i^d \end{bmatrix}^T$: desired position
- $p_c = \begin{bmatrix} x_c & y_c \end{bmatrix}^T$: position of virtual leader
- $p_{iF}^d = \begin{bmatrix} x_{iF}^d & y_{iF}^d \end{bmatrix}^T$: desired deviation of the ith robot

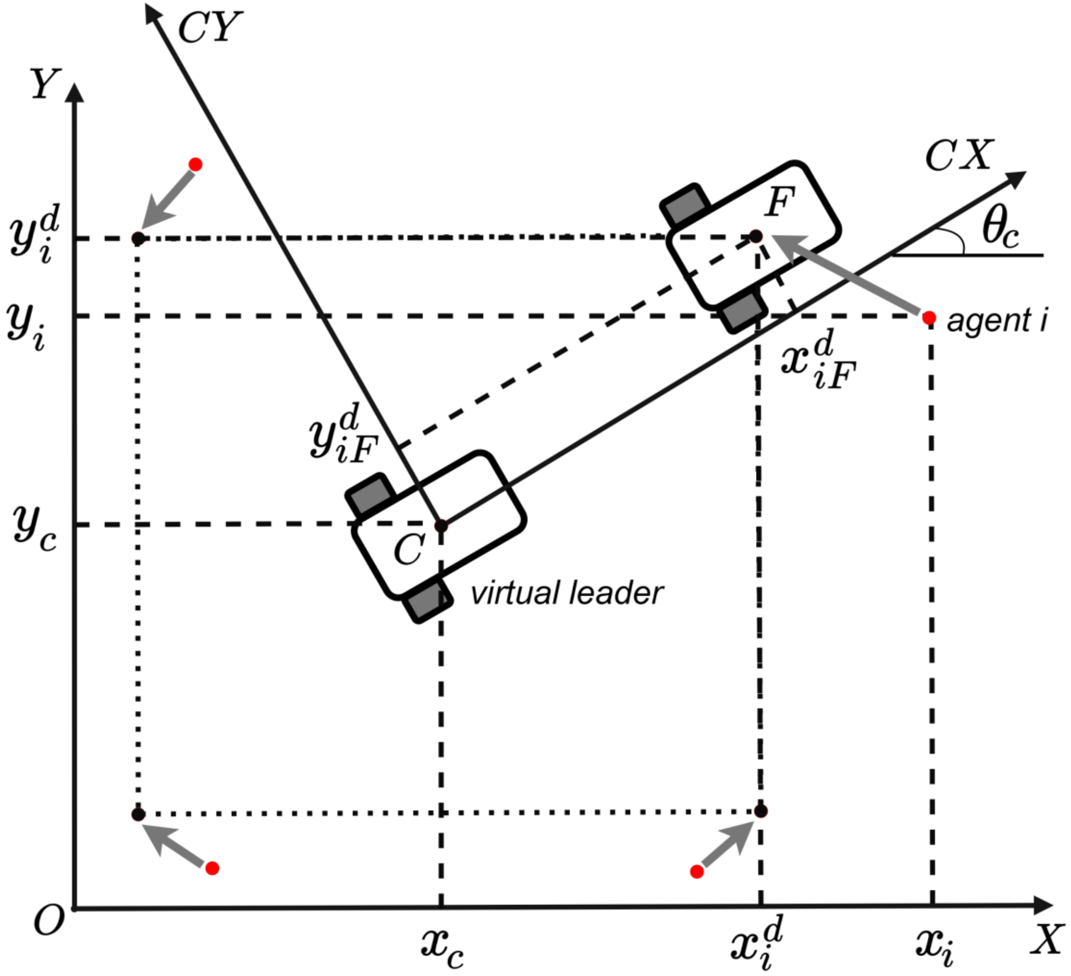
Where $x_{iF}^d = l_i \cos(\phi_i)$, $y_{iF}^d = l_i \sin(\phi_i)$

+ l_i : the distance from the formation center to each agent

+ ϕ_i : is the angular offsets for each agent

$$\phi_i = \frac{\pi}{2}(i-1)\,,\ i = 1, \dots, 4 \text{ for square shape}$$

$$\phi_i = \frac{2\pi}{3}(i-1)\,,\ i = 1, \dots, 3 \text{ for triangle shape}$$



Square formation with a known virtual leader

➔ Control law (32) using consensus algorithm can be applied.

ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

C. Tracking control design:

1. Control strategy

The consensus-based control law requires each WMR to approximate the double-integrator dynamics (30):

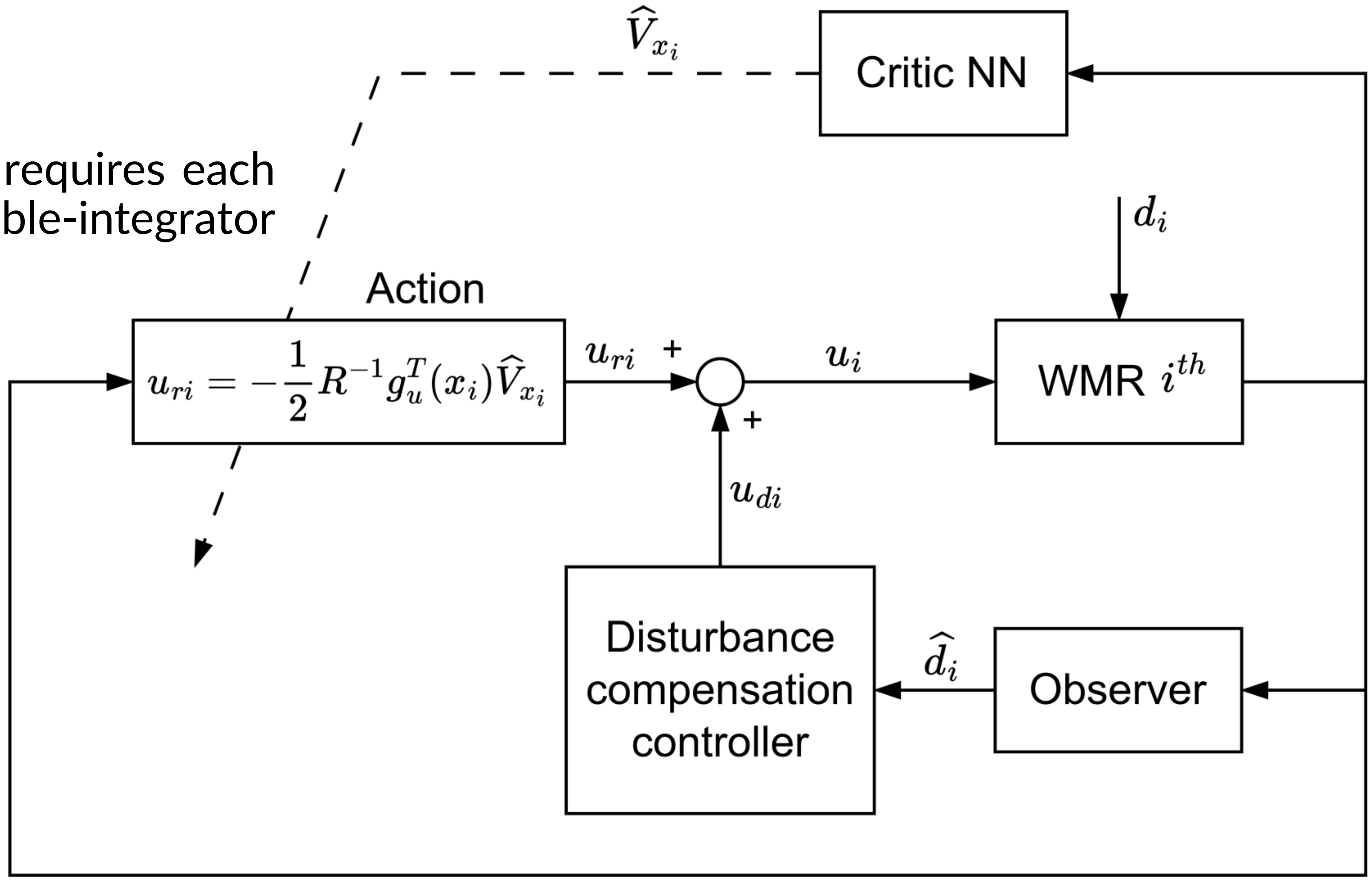
$$\begin{cases} \dot{p}_{ri} = v_{ri} \\ \dot{v}_{ri} = u_i \end{cases}$$

The kinematics of the i^{th} robot under disturbance is obtained from (5):

$$\dot{p}_i = S v_i \quad (40)$$

Define the tracking error:

$$e_i = p_i - p_{ri} \quad (41)$$



ADP control structure combined with disturbance observer

4. Multi-agent control using ADP approach

C. Tracking control design:

1. Control strategy

The error dynamics is derived as:

$$\dot{e}_i = \dot{p}_i - \dot{p}_{ri} = Sv_i + d_p - v_{ri} \qquad (42)$$

The time derivative of (42): $\ddot{e}_i = \dot{S}v_i + S\dot{v}_i + \dot{d}_p - u_i \qquad (43)$

The dynamic model for i^{th} robot (23) can be rewritten as:

$$\dot{v}_i = -M^{-1}Bv_i + M^{-1}\tau_i + d_v \quad (44)\text{ , where } \; d_v = -M^{-1}(Q\ddot{\gamma} + C\dot{\eta} + G\ddot{\eta} + \tau_d)$$

Substitute (44) into (43) : $\ddot{e}_i = -SM^{-1}Bv_i + u_i^{new} + \dot{S}v_i + Sd_v + \dot{d}_p \quad (45)$

Where $u_i^{new} = \mathrm{S}M^{-1}\tau_i - u_i$ is the transformed control input.

The control torque for the i^{th} WMR is calculated as:

$$\boxed{\tau_i = MS^{-1}(u_i^{new} + u_i) \quad (46)}$$

4. Multi-agent control using ADP approach

C. Tracking control design:

1. Control strategy

Define $x_{1i} = e_i$, $x_{2i} = \dot{x}_{1i} + \lambda x_{1i}$, the time derivative of x_{1i} and x_{2i} is obtained as:

$$\begin{cases} \dot{x}_{1i} = \dot{e}_i = Sv_i + d_p - v_{ri} \\ \dot{x}_{2i} = Ex_{2i} - E\lambda x_{1i} + u_i^{new} + d_i \end{cases} \quad (47)$$

Where: $E = E_1 S^{-1}$, $E_1 = -SM^{-1}B$, $d_i = \xi_{2i} - E\xi_{1i}$,

$$\xi_{1i} = d_p - v_{ri} \;,\; \xi_{2i} = Sd_v + \dot{S}v_i + \dot{d}_p + \lambda Sv_i + \lambda\xi_{1i}$$

Rewrite (47) in state space form: $\dot{x}_i = f(x_i) + g_u u_i^{new} + g_d d_i \quad (48)$

Where $x_i = \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix}$, $f(x_i) = \begin{bmatrix} x_{2i} - \lambda x_{1i} \\ Ex_{2i} - \lambda Ex_{1i} \end{bmatrix}$, $g_u = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $g_d = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

The control input is derived as follows to achieve optimal performance :

$$\boxed{u_i^{new} = u_{ri}(x_i) + u_{di}(x_i) \quad (49)}$$

C. Tracking control design:

2. Observer-based controller design

A disturbance observer is used to estimate disturbance d_i as follows:

$$\begin{cases} \hat{d}_i = z_i + \rho(x_i) \\ \dot{z}_i = -h(x_i)\{g_d(x_i)[z_i + \rho(x_i)] + f(x_i) + g_u(x_i)u_i^{new}\} \end{cases} \qquad (50)$$

Where:

- $\hat{d}_i$: the estimation of d_i
- z_i : the internal state of the observer
- $\rho(x_i)$: the nonlinear function vector to be designed
- $h(x_i)$: the gain of the observer, defined by $h(x_i) = \frac{\partial \rho(x_i)}{\partial x_i}$

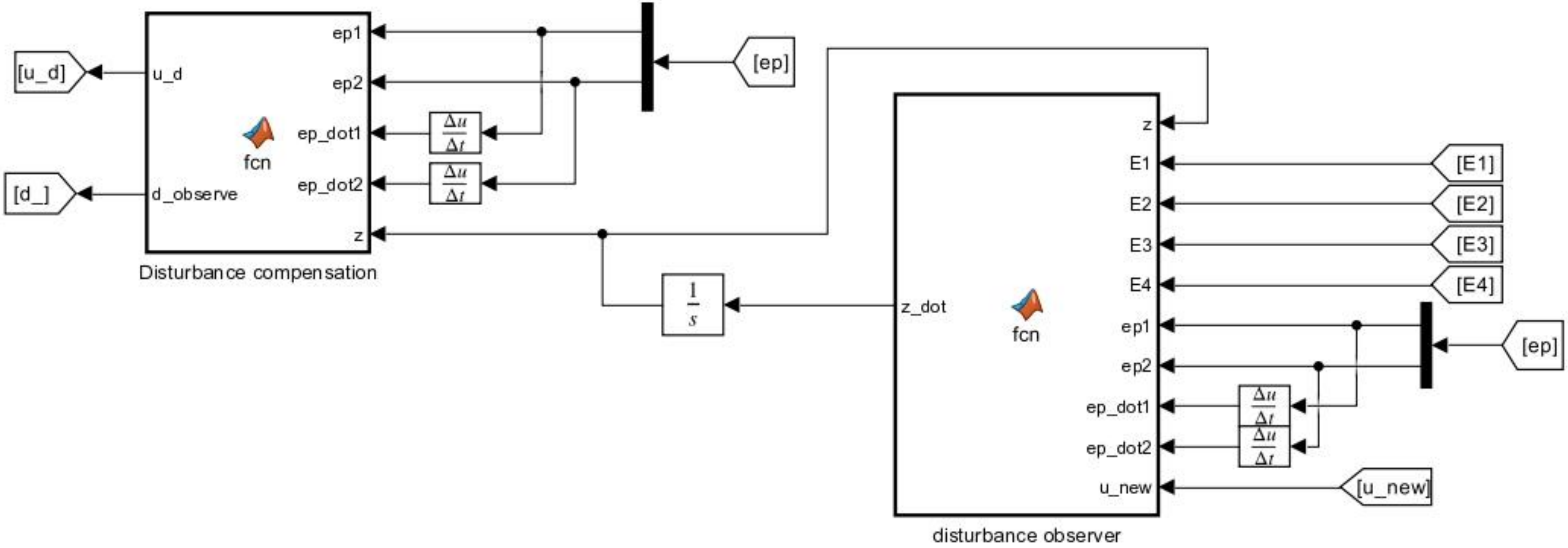
The compensation control component is derived as:

$$\boxed{u_{di}(x_i) = -\hat{d}_i \qquad (51)}$$

C. Tracking control design:

2. Observer-based controller design

The observer-based controller simulation implementation:



C. Tracking control design:

3. Adaptive optimal controller design

In the case of $d_i = 0$ the error system (48) is rewritten as:

$$\dot{x}_i = f(x_i) + g_u u_i^{new} \qquad (52)$$

Define the following cost function:

$$J(x_i) = \int_t^{\infty} r\left(x_i(\tau_i), u_{ri}\big(x_i(\tau_i)\big)\right) d\tau_i \qquad (53)$$

Where the utility function is defined as: $r = x_i^T(\tau_i) Q x_i(\tau_i) + u_{ri}^T(x_i) R u_{ri}(x_i)$

Define the Hamiltonian function of the problem as:

$$H = \left(\frac{\partial J}{\partial x_i}\right)^T + r(x_i, u_{ri}) = \left(\frac{\partial J}{\partial x_i}\right)^T + x_i^T Q x_i + u_{ri}^T R u_{ri} \qquad (54)$$

C. Tracking control design:

3. Adaptive optimal controller design

For the system (52) to have the optimal solution, there must exist an optimal cost function $V(x_i, u_{ri})$ defined by:

$$V(x_i, u_{ri}) = \min_{u_{ri}}\big(J(x_i)\big) \qquad (55)$$

Which satisfies the HJB equation:

$$0 = \min_{u_{ri}}\big(H(x_i, u_{ri}, V)\big) = \left(\frac{\partial V}{\partial x_i}\right)^T \dot{x}_i + x_i^T Q x_i + u_{ri}^T R u_{ri} \qquad (56)$$

The optimal control component u_{ri} can be derived as:

$$u_{ri} = \underset{u_{ri}}{\operatorname{argmin}}\big(H(x_i, u_{ri}, V)\big) \qquad (57)$$

Solve (57) using (56) obtains:

$$\boxed{u_{ri} = -\frac{1}{2} R^{-1} g_u^T \frac{\partial V}{\partial x_i} \qquad (58)}$$

C. Tracking control design:

3. Adaptive optimal controller design

The value function $V(x_i, u_{ri})$ is approximated by a NN as:

$$V(x_i, u_{ri}) = W_i^T \phi(x_i) + \epsilon(x_i) \quad (59)$$

The estimated cost function:

$$\hat{V}\,(x_i, u_{ri}) = \widehat{W}_i^T \phi(x_i) \qquad (60)$$

The approximated Hamiltonian function:

$$H\big(x_i, u_{ri}, \widehat{W}_i\big) = \widehat{W}_i^T \left(\frac{\partial \phi}{\partial x_i}\right) \dot{x}_i + x_i^T Q x_i + u_{ri}^T R u_{ri} = e_i \quad (61)$$

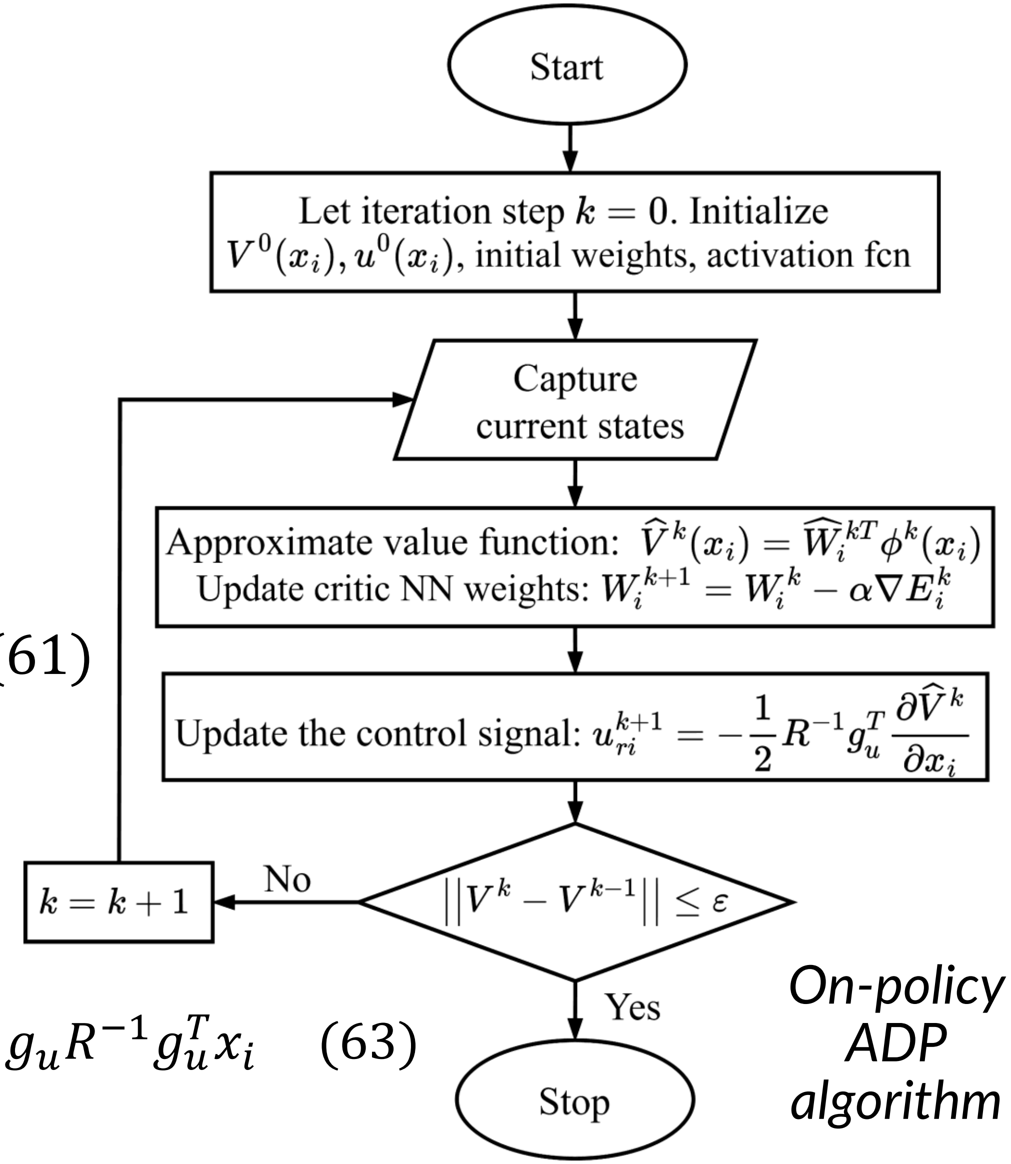
Update $\widehat{W}_i$ to minimize the squared residual error:

$$E_i = \frac{1}{2} e_i^T e_i \quad (62)$$

The weight update law is given by:

$$\dot{\widehat{W}}_i = -\alpha_1 \frac{\hat{\sigma}_i}{\left(\hat{\sigma}_i^T \hat{\sigma}_i + 1\right)^2} \left(\hat{\sigma}_i^T \widehat{W}_i + Q(x_i) + u_{ri}^T R u_{ri}\right) + \frac{1}{2} \alpha_2 \frac{\partial \phi}{\partial x_i} g_u R^{-1} g_u^T x_i \quad (63)$$

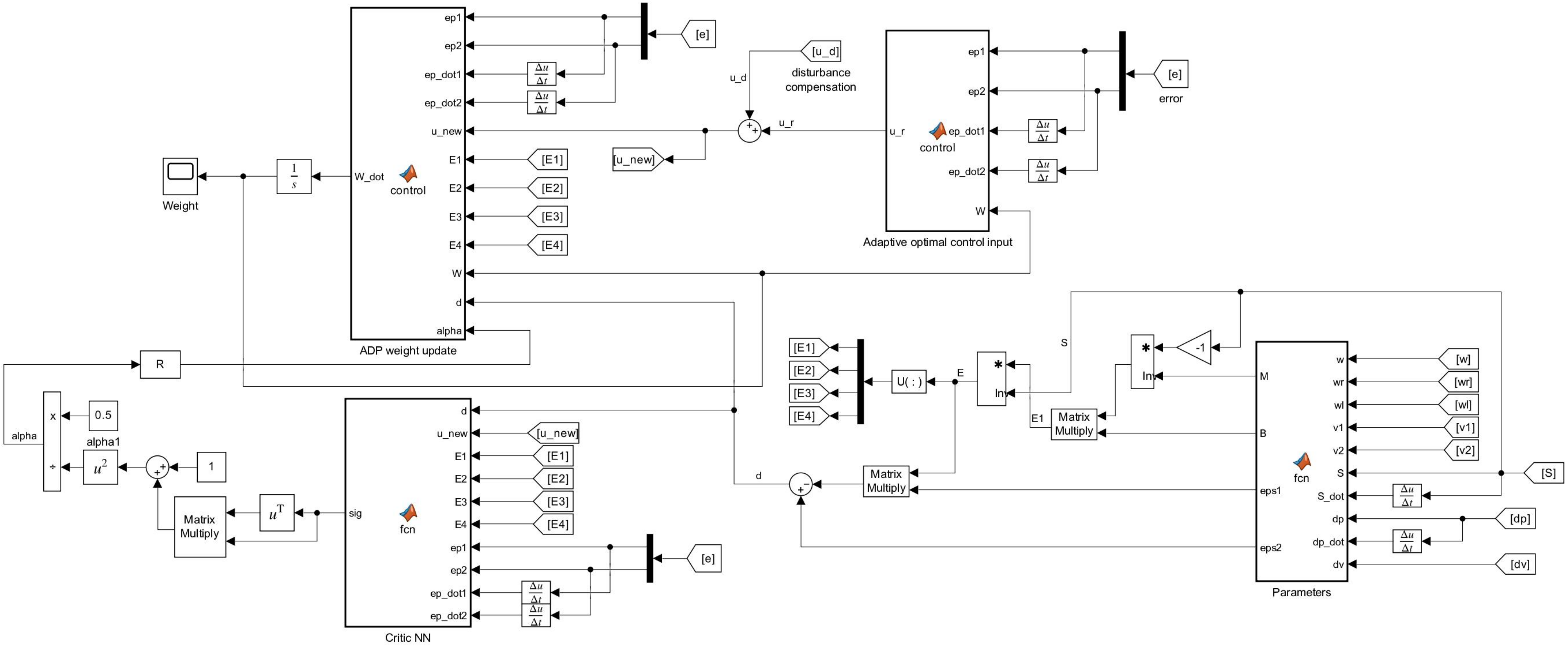
Where $\hat{\sigma}_i = (\partial \phi / \partial x_i) \dot{x}_i$



On-policy ADP algorithm

4. Multi-agent control using ADP approach

The adaptive optimal control simulation implementation:

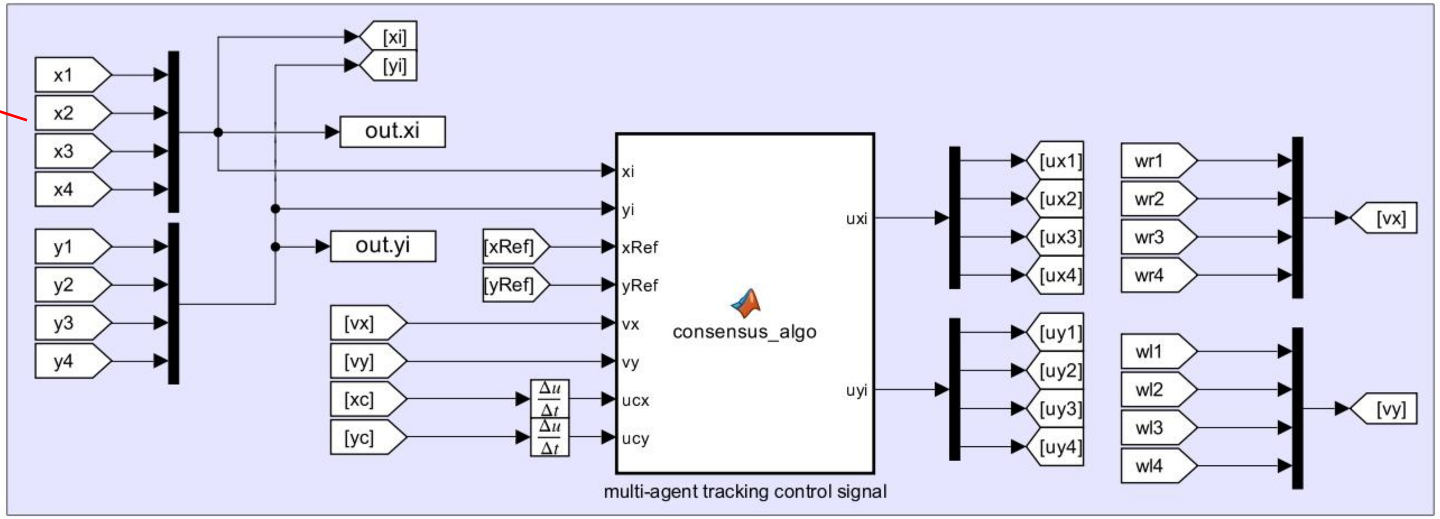


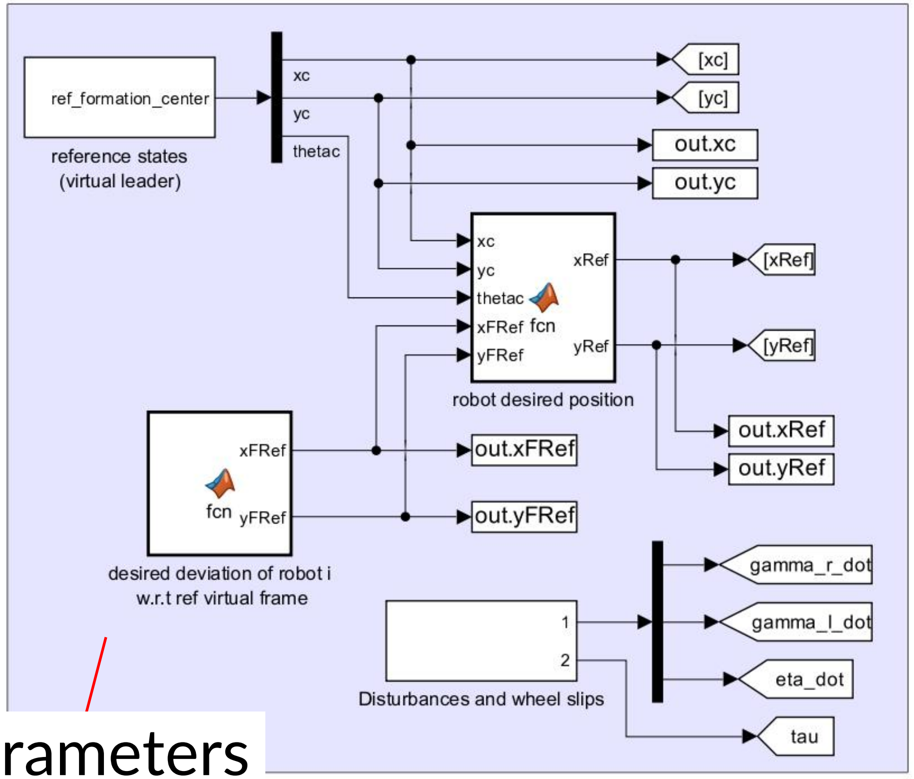
Simulation and Results

5. Simulation implementation

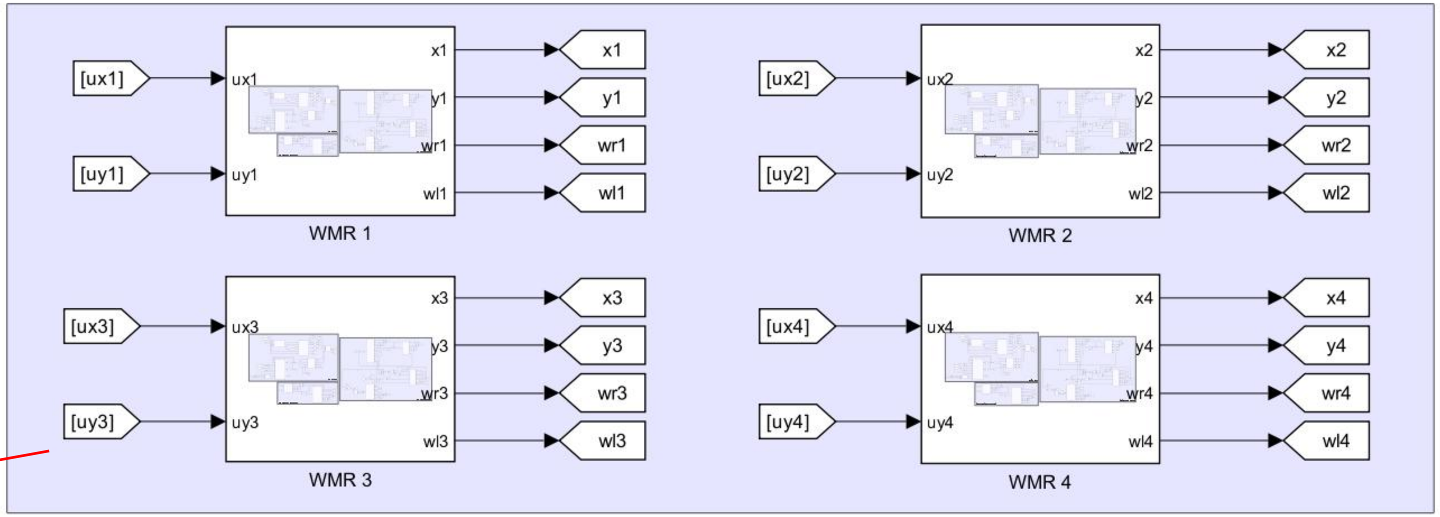
Overall system simulation :

Multi-agent control





Parameters generation



WMR system

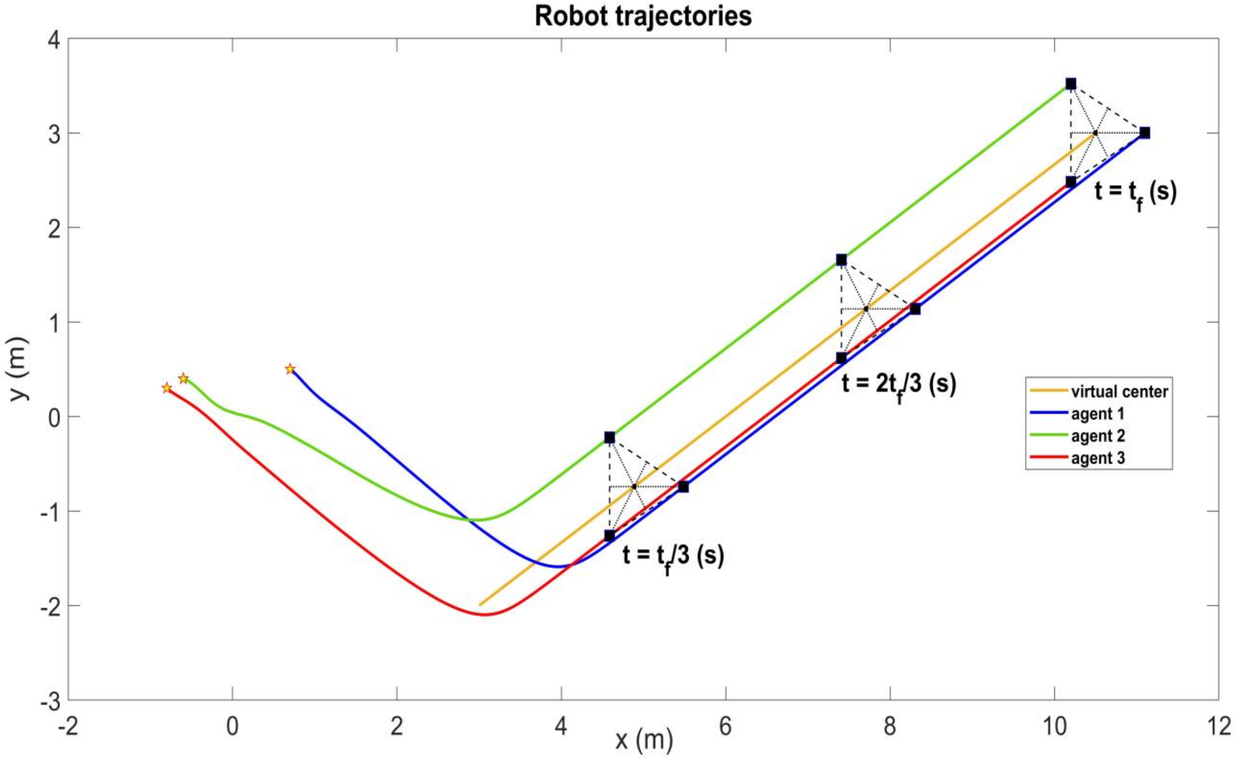
5. Simulation results

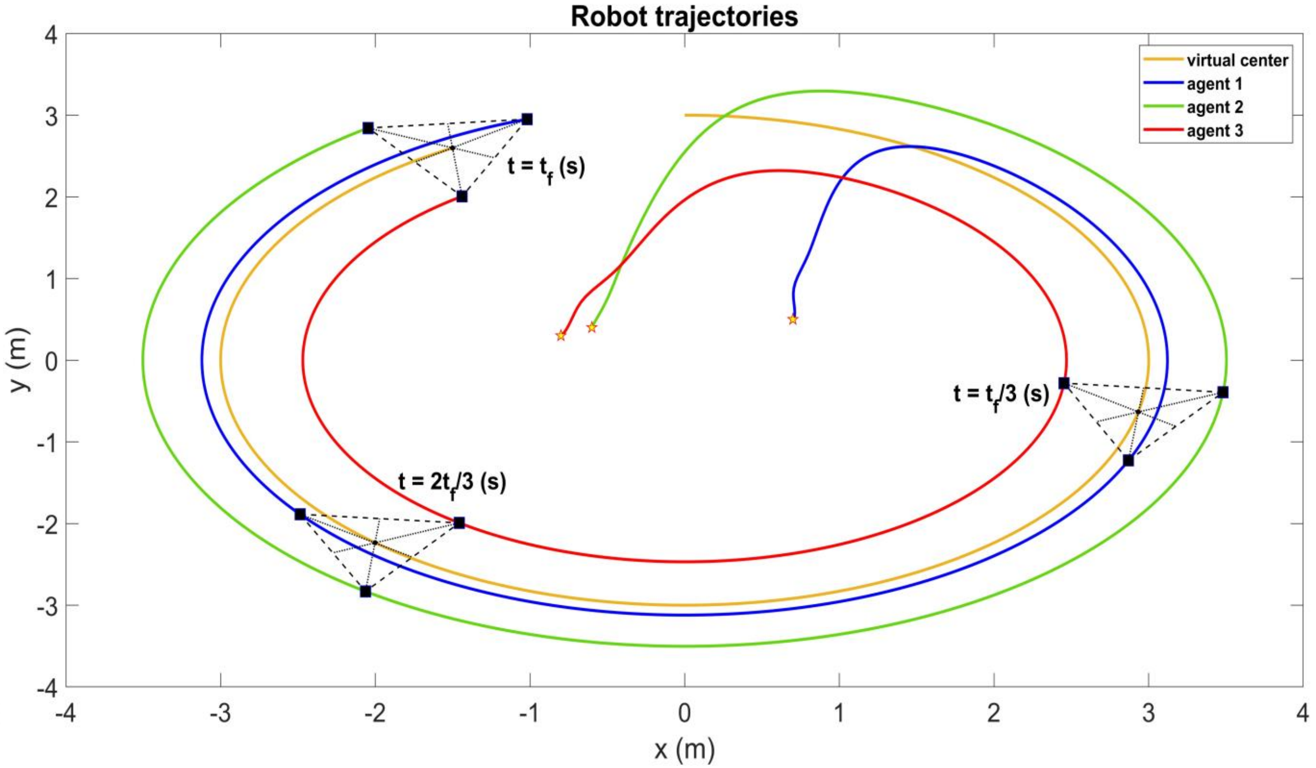
A. Formation tracking results of multiple WMRs

- Scenario 1: Triangle formation

Line trajectory: $\begin{cases} x_c = 3 + 0.75t \\ y_c = -2 + 0.5t \end{cases}$

Circle trajectory: $\begin{cases} x_c = 3\sin\left(\frac{\pi}{6}t\right) \\ y_c = 3\cos\left(\frac{\pi}{6}t\right) \end{cases}$

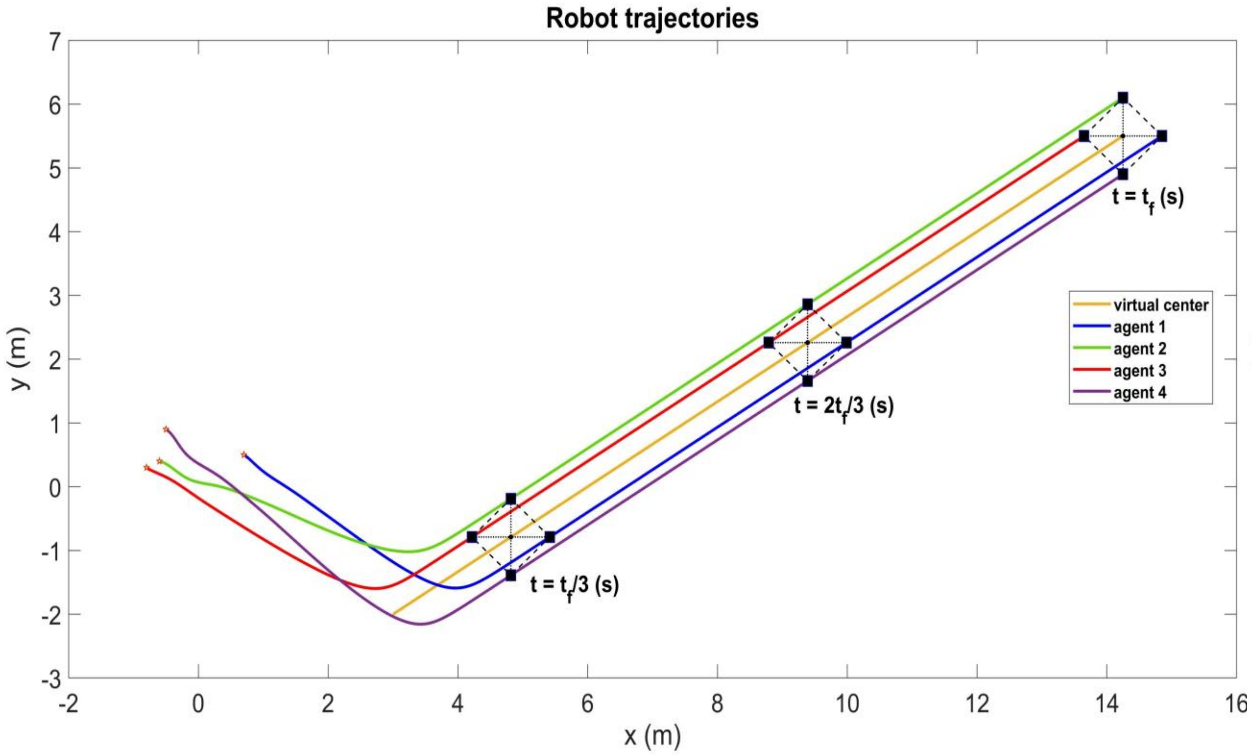




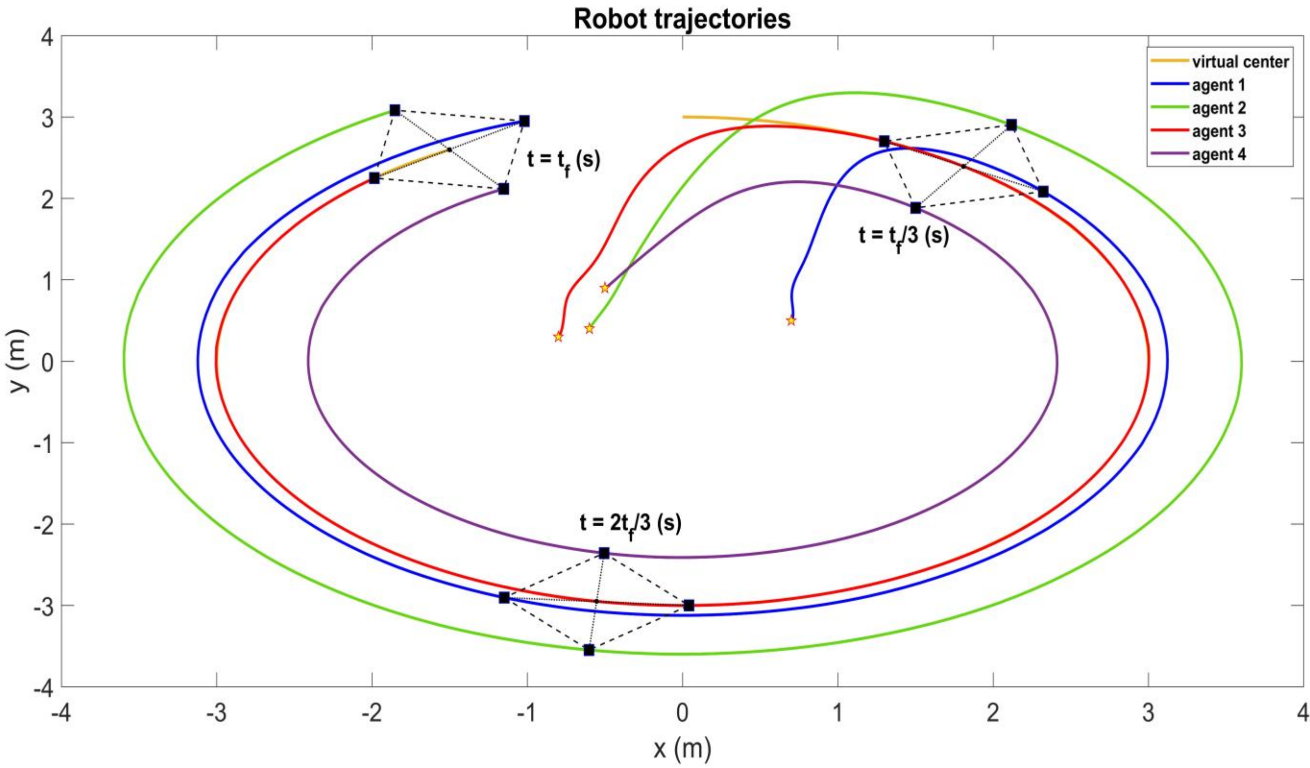
A. Formation tracking results of multiple WMRs

- Scenario 2: Square formation

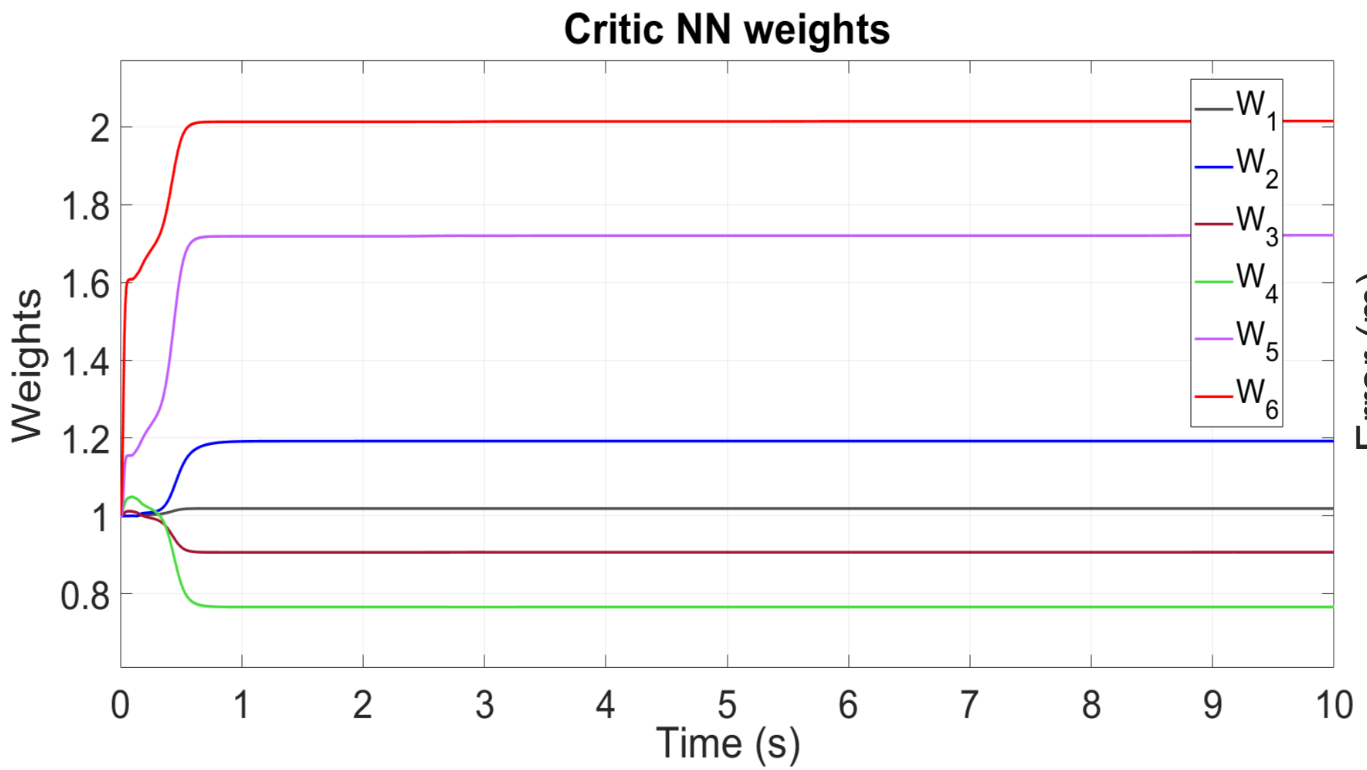
Line trajectory

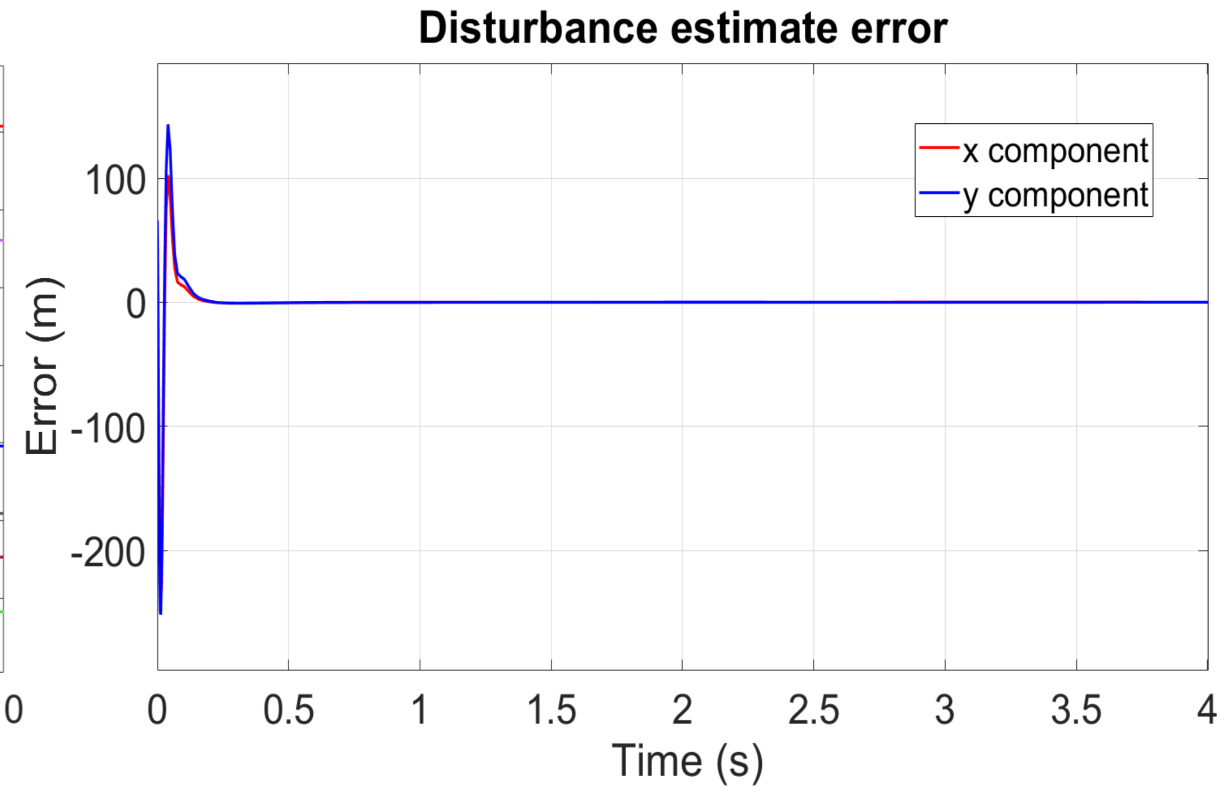


Circle trajectory



A. Formation tracking results of multiple WMRs



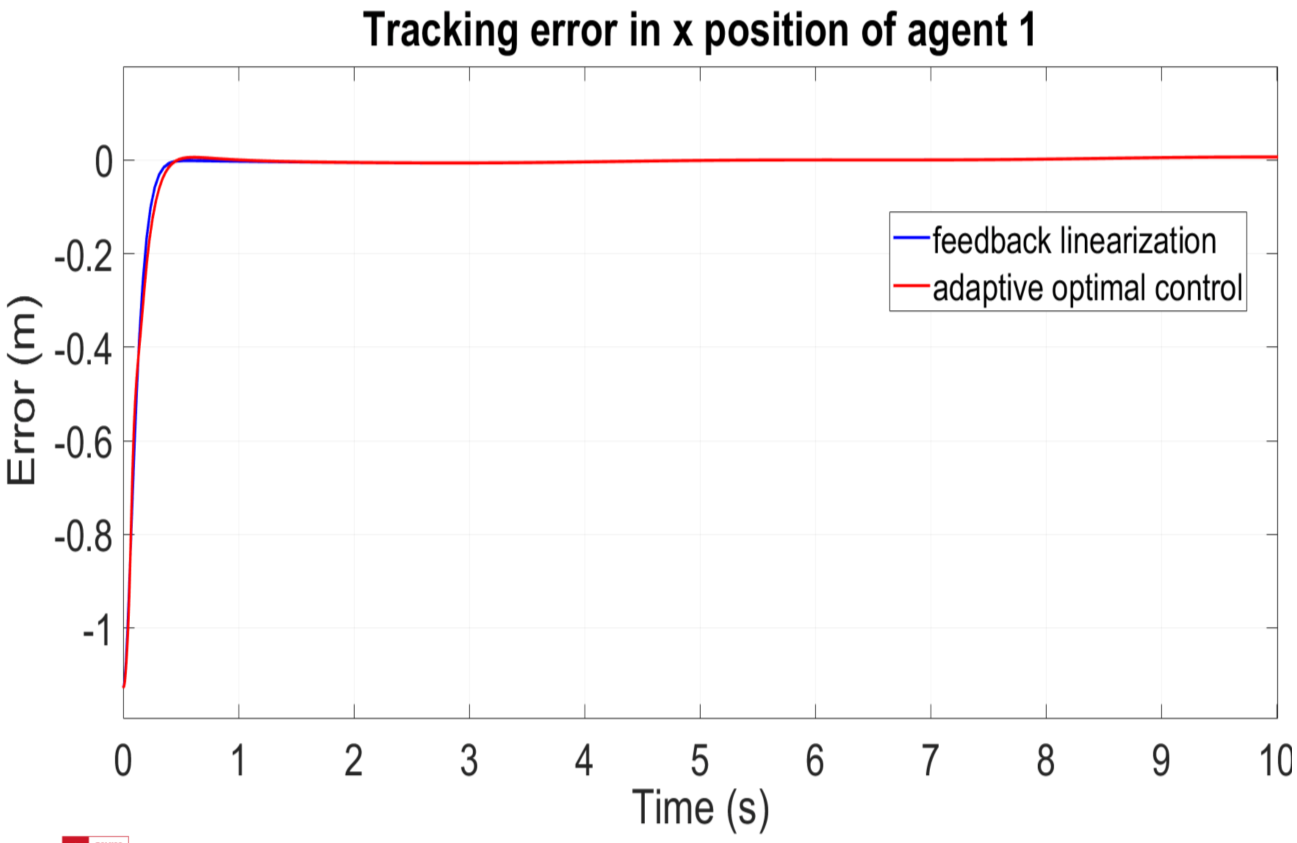


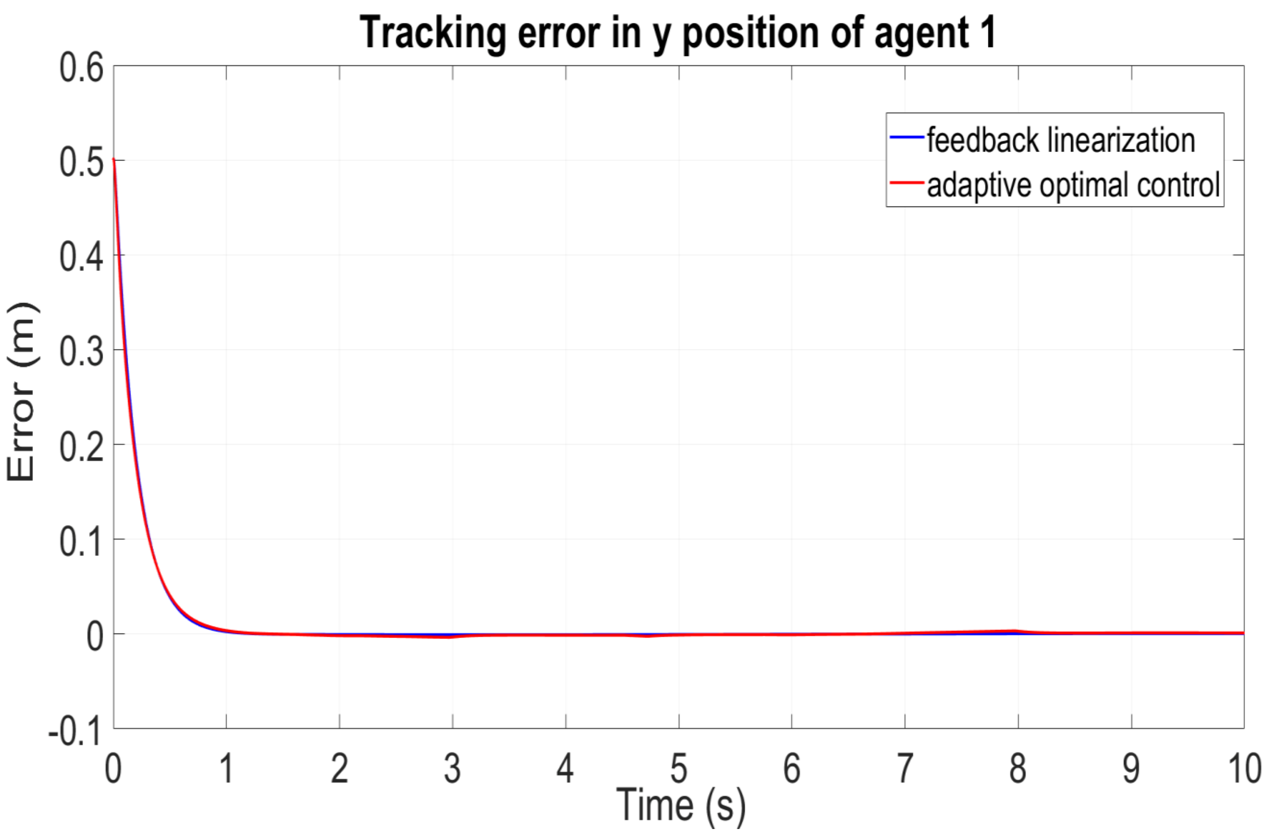
Results of ADP-based optimal control for agent 1 tracking triangle formation while following line trajectory

B. Comparison between ADP-based optimal control method and conventional feedback linearization control method:

In case of no disturbance:

Curved trajectory: $\begin{cases} x_c = -t \\ y_c = \sin(0.5t) + 0.5t + 1 \end{cases}$

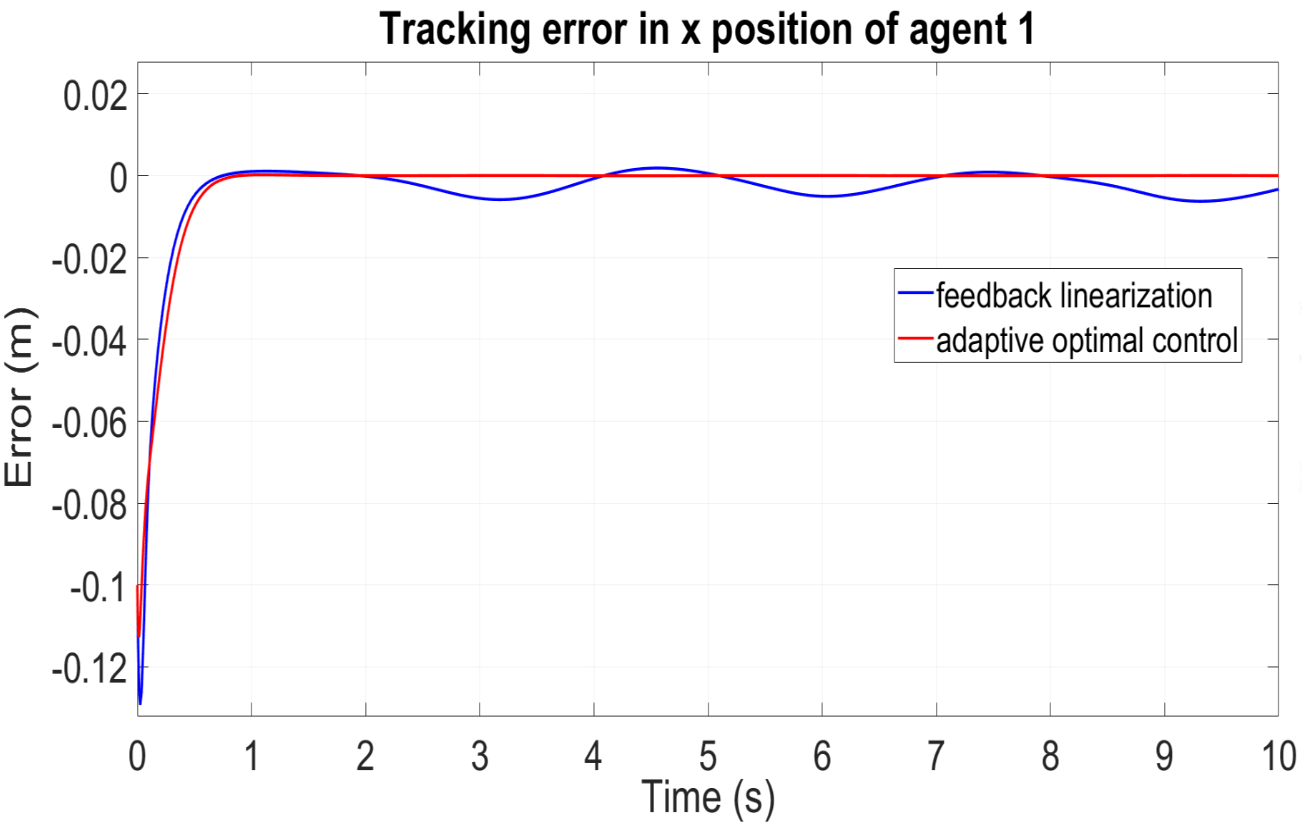


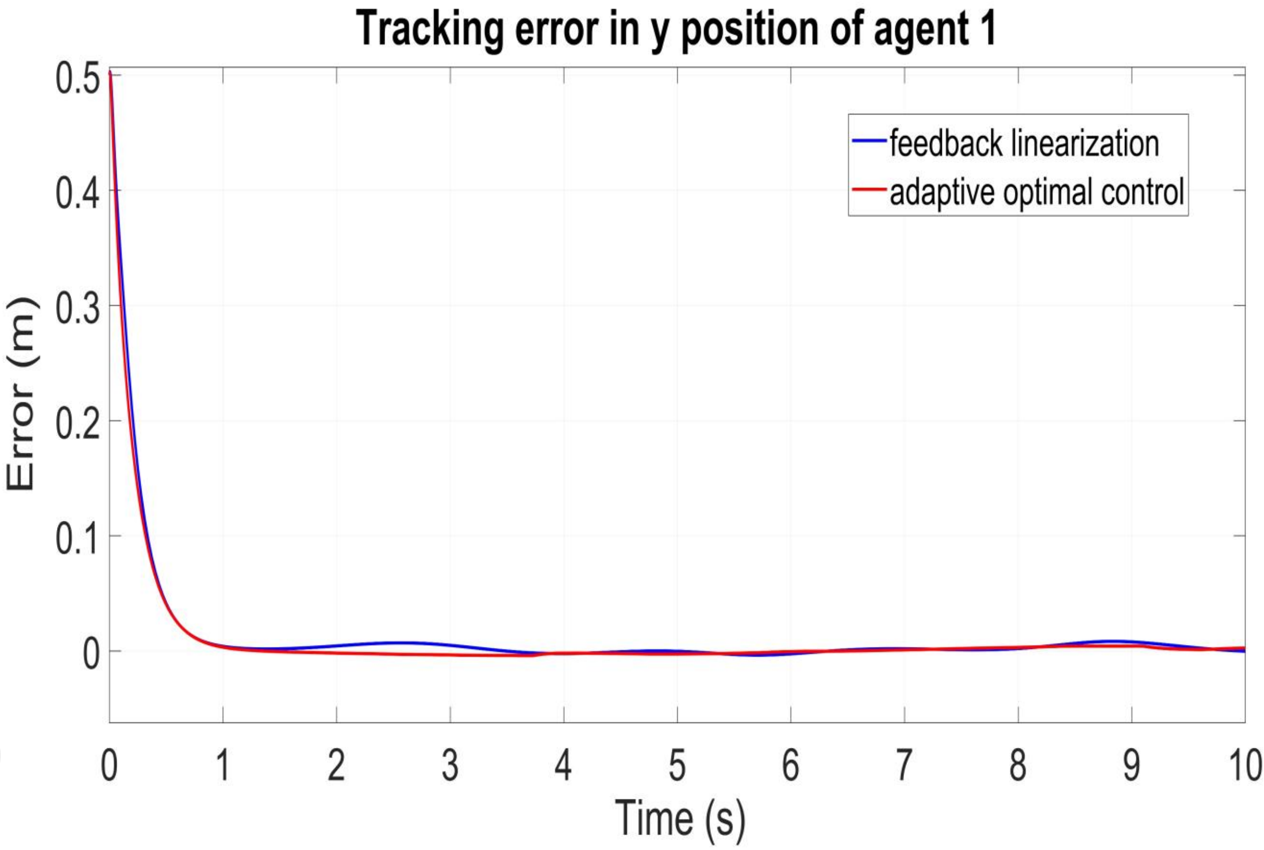


5. Simulation results

B. Comparison between ADP-based optimal control method and conventional feedback linearization control method:

In case of wheel slip and external disturbance:





6. Conclusion

- In general, the WMR system achieves the desired formations while following the predefined trajectories.
- The approach assumes each agent has consistent understanding of the formation center. Future study can focus on designing consensus algorithm to drive each agent’s instantiation of the virtual center to a common value.
- The proposed method requires knowledge of the system’s internal dynamics. Future improvement may focus on developing model-free control algorithms that rely solely on system state data.
- Possibly develop more effective multi-agent control algorithms. These advancements may include distance-based control, direction vector-based control, and graph rigidity theory-based methods.

Reference

[1] N. Van Tinh et al., “Neural network-based adaptive tracking control for a nonholonomic wheeled mobile robot subject to unknown wheel slips,” Journal of Computer Science and Cybernetics, vol. 33, no. 1, pp. 70–85, 2017.

[2] N. M. H. Trịnh Hoàng Minh, “Điều khiển hệ đa tác tử,” 2021.

[3] W. Ren and R. W. Beard, Distributed consensus in multi-vehicle cooperative control, vol. 27. Springer, 2008.

[4] W.-H. Chen, “Disturbance observer based control for nonlinear systems,” IEEE/ASME transactions on mechatronics, vol. 9, no. 4, pp. 706–710, 2004.

[5] H. Van Doan and N. T.-T. Vu, “Robust optimal control for uncertain wheeled mobile robot based on reinforcement learning: Adp approach,” Bulletin of Electrical Engineering and Informatics, vol. 13, no. 3, pp. 1524–1534, 2024.

[6] R. Song and F. L. Lewis, “Robust optimal control for a class of nonlinear systems with unknown disturbances based on disturbance observer and policy iteration,” Neurocomputing, vol. 390, pp. 185–195, 2020

HUST

hust.edu.vn fb.com/dhbkhn

THANK YOU !