ISA Transactions 109 (2021) 89–101

Contents lists available at ScienceDirect

ISA Transactions

journal homepage: www.elsevier.com/locate/isatrans

Research article

Neural network-based model predictive tracking control of an uncertain robotic manipulator with input constraints

Erlong Kang [a,b,c], Hong Qiao [a,b,d,*], Jie Gao [a,b,c], Wenjing Yang [e]

[a] The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China
[c] Beijing Key Laboratory of Research and Application for Robotic Intelligence of Hand-Eye-Brain Interaction, Beijing 100190, China
[d] CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai 200031, China
[e] State Key Laboratory of High Performance Computing, National University of Defense Technology, Changsha 410073, China

ARTICLE INFO

Article history:
Received 4 December 2019
Received in revised form 9 July 2020
Accepted 3 October 2020
Available online 8 October 2020

Keywords:
Model predictive control
Neural network
Robotic manipulator
Unknown dynamics
Online learning estimation
Input constraints

ABSTRACT

This paper proposes a neural network-based model predictive control (MPC) method for robotic manipulators with model uncertainty and input constraints. In the presented NN-based MPC structure, two groups of radial basis function neural networks (RBFNNs) are considered for online model estimation and effective optimization. The first group of RBFNNs is introduced as a predictive model for the robotic system with online learning strategies for handling the system uncertainty and improving the model estimation accuracy. The second one is developed for solving the optimization problem. By taking into account an actor–critic scheme with different weights and the same activation function, adaptive learning strategies are established for balancing between optimal tracking performance and predictive system stability. In addition, aiming at guaranteeing the input constraints, a nonquadratic cost function is adopted for the NN-based MPC. The ultimately uniformly boundedness (UUB) of all variables is verified through the Lyapunov approach. Simulation studies are conducted to explain the effectiveness of the proposed method.

1. Introduction

In recent years, the robotic control design has been getting sustained attention from both industry and academia. Many control theories, such as neural network control, fuzzy control, sliding mode control and other control methods [1–7] have been successfully applied into robotic systems and related systems, independently or in combination. With the continuous expansion of robotic applications during the past decades, the optimal control performance has been receiving more and more attention in addition to system stability. Furthermore, the model uncertainty and input constraints are also challenges for the control design of an actual robotic system. It is therefore crucial to design an effective control strategy for robotic manipulators, which can balance between optimal control performance and system stability, compensate for the effect of model uncertainty, and satisfy the input constraints.

Model predictive control (MPC), also named receding horizon control, is a powerful optimal control strategy. MPC has several attractive characteristics, for example, it deals with multivariable and constrained control problems [8]. Until now it has been successfully applied to the process industry [9], power electronics industry [10], smart energy systems [11], motors control for electric vehicles [12] and robotic systems, especially mobile robots [13–16].

Two key issues need to be studied for solving robotic control problems with MPC. One lies in realizing robustness against model uncertainties, the other lies in effective optimization based on the predictive model. For the first but challenging issue, many significant results have been investigated centered on nominal systems with disturbance. The nominal dynamics are utilized as the predictive model for MPC. For known or partially known systems, the known dynamics are adopted as the nominal model [15–20]. The disturbance is handled by robust MPC [16,17], tube MPC [18,19], min–max MPC [20], etc., or is compensated by an extra robust controller [21]. In [21], a linear MPC with an integral sliding mode (ISM) controller is studied for robotic manipulators. Partially known robotic dynamics are used as a nominal model, feedback linearization is used to transform the nonlinear problem into a linear form, and the ISM controller is used to compensate disturbance and unknown dynamics. In [22], a path-following MPC strategy based on known dynamics is

* Corresponding author at: The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China.
E-mail address: hong.qiao@ia.ac.cn (H. Qiao).

https://doi.org/10.1016/j.isatra.2020.10.009
0019-0578/© 2020 ISA. Published by Elsevier Ltd. All rights reserved.

proposed for an industrial robot. However, these methods require a clear nominal model, which may not be suitable for a robot with completely unknown dynamics. For unknown systems, neural network (NN) model [23–29], fuzzy model [30], Gaussian process model [31], etc. are utilized as nominal dynamics after appropriate off-line training. In [30], MPC based on a Takagi–Sugeno (T–S) fuzzy model is adopted for 2-DOF robotic arms. In [31], a Gaussian process MPC scheme is developed for the robotic arm, in which a Gaussian process based on off-line data is adopted as a nominal model, and an extended Kalman filter-based observer is used to compensate residual disturbance. However, these methods do not think about model uncertainty online.

Some researchers also focus on combining adaptive NN with MPC. Wang et al. [9] propose a double-layers architecture controller, in which adaptive NN is used for the lower layer to approximate the unknown dynamics. But the two-layer-structure MPC is appropriate for the industrial process rather than a robotic manipulator. Wu et al. [32] develop an adaptive MPC for motor system, a two-layer recursive NN with extended-Kalman-filter-based parameter learning is used for speed predictor. But the stability analysis is not been considered. Chen et al. [33] present a tube-based MPC for nonholonomic mobile robots. An adaptive NN controller with disturbance observer is used for unknown dynamics, independently of MPC strategy for kinematic constraints, which is unsuitable for robotic manipulators. Farrokhi et al. [34] introduce an adaptive nonlinear MPC for hybrid position/velocity control of robot manipulators. But off-line training should be considered to avoid irrational control signals at the beginning of an operation. Therefore, for the first key issue, the MPC strategy for robotic manipulators needs to be further developed subject to handle the model uncertainty online.

For the second key issue, several approaches are proposed for the optimization of MPC. In [15,21,35], linearization models are developed for nonlinear systems, efficient methods such as linear quadratic regulator (LQR) and linear matrix inequalities (LMI) are adopted. In [16,17], the event-trigger mechanism is utilized with MPC for reducing the computational burden of MPC. In [36,37], intelligent algorithms such as genetic algorithm and particle swarm optimization are used for solving the optimization problem of MPC. In [38,39], NN solvers based on neurodynamic optimization are proposed for solving MPC. In [23,27], ADP-based methods are studied, where critic and actor NN are constructed for estimating cost function and input signal, respectively. However, there is little research located in effective optimization based on the online estimating predictive model.

According to previous discussions, designing a suitable MPC strategy for robotic manipulators, which estimates the unknown dynamics online and balances between the optimal control performance and system stability, is still an unsolved problem. The difficulties lie in estimating the unknown model online, solving the optimization problem based on the online-updating predictive model and ensuring the stability of the whole system under above conditions. In this paper, we develop an NN-based MPC strategy for robotic manipulators with unknown plant model and input constraints. The main contributions of this paper are summarized as follows:

(1) An NN-based MPC structure containing two groups of NNs is proposed for robotic manipulators. The first group of NNs used as the predictive model for MPC is established to estimate unknown robotic dynamics. Online updating laws of NNs' weights are proposed without requiring knowledge of the system. Furthermore, it is proved that the estimation error is UUB according to the Lyapunov theorem.

(2) The second group of NNs, which adopts the actor–critic scheme with the same activation function but different weights, is established for solving the optimization problem of MPC. An adaptive learning approach based on the Hamiltonian function is built to guarantee the optimal control performance and predictive system stability. Meanwhile, the input constraints are guaranteed by employing a suitable integrand function of input signals in the cost function of MPC.

(3) The stability of the closed-loop system is proved using the Lyapunov theorem and mathematical induction (MI). All variables remain UUB under the developed control strategy.

The rest of this paper is organized as follows. In Section 2, some preliminaries used later and robotic system dynamics are introduced. Section 3 illustrates the main results of this paper, concluding the establishing of NN-based MPC strategy and stability analysis. The performance of the proposed control strategy is shown in Section 4 by co-simulation based on CoppeliaSim (V-REP) and Matlab, and conclusions of this paper are given in the last section.

2. Preliminaries and problem formulation

2.1. Preliminaries

Lemma 1 (*[40] First Mean Value Theorem for Integrals*)**.** *Let $f(x)$ is continuous on $[a, b]$, $g(x)$ is integrable and sign-invariant on $[a, b]$. Then there exists $\varepsilon \in [a, b]$, such that*

$$\int_a^b f(x)\, g(x)\, dx = f(\varepsilon) \int_a^b g(x)\, dx \tag{1}$$

Lemma 2. *Let $f(x)$ is continuous on $[a, b]$, then we have*

$$\int_a^b f(x)\, dx = \sum_{i=1}^{N-1} \frac{1}{2} \left(f(x_i) + f(x_{i+1})\right)(x_{i+1} - x_i) + \varsigma \tag{2}$$

where $x_1 = a, x_N = b$, ς is the integral error which is bounded, i.e. there exists $\varsigma_0 > 0$, such that $\|\varsigma\| \le \varsigma_0$.

Lemma 3 (*[41]*)**.** *Let $A \in \Re^{n\times n}$ be a semi-definite symmetric matrix, then all the eigenvalues of A are real and nonnegative. $\forall x \in \Re^n$, there exists $\underline{\lambda}_A \|x\|^2 \le x^T A x \le \bar{\lambda}_A \|x\|^2$, where $\underline{\lambda}_A \geqslant 0$, $\bar{\lambda}_A$ are the minimum and maximum eigenvalue of A, respectively, $\|.\|$ represents the standard Euclidean norm.*

Lemma 4 (*[5,42]*)**.** *Let Lyapunov function $V(x(t))$ be a continuous and positive definite function, with bounded initial value $V(x(0))$. If the inequality $\dot{V}(x) \le -c_1 V(x) + c_2$ holds, where c_1 and c_2 are positive constants, then $V(x(t))$ is bounded. Furthermore, the solution $x(t)$ of the underlying system is uniformly bounded.*

2.2. Problem formulation

Consider an n-link robotic manipulator formulated by the following dynamics [2,21,41]:

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = \tau \tag{3}$$

where q, $\dot{q}$ and $\ddot{q} \in \Re^n$ represent the joint position, velocity and acceleration vectors, respectively. $M(q) \in \Re^{n\times n}$ denotes a symmetric positive definite inertia matrix, $C(q, \dot{q})\dot{q} \in \Re^n$ represents the Centripetal and Coriolis force, $G(q) \in \Re^n$ represents the gravitational force, $\tau \in \Re^n$ represents the input torque. Generally speaking, the input torque of the robot is bounded, which must be considered while designing the control strategy. In this paper, the input constraints are expressed by

$$|\tau_i(t)| \le \lambda, i = 1, 2, \ldots, n \tag{4}$$

Property 1 (*[2]*)**.** *The inertia matrix $M(q)$ is symmetric and positive definite.*

The control objective is to design a suitable control strategy that satisfies the constraints (4), such that the system variable q can track a given desired trajectory $q_d(t) = [q_{d1}(t), q_{d2}(t), \dots, q_{dn}(t)]^T$, while trading off between the tracking performance and the stability of the closed-loop system.

Assumption 1. The desired trajectory q_d is bounded, smooth and twice differentiable. Therefore, there exist $d > 0$, $d_1 > 0$, $d_2 > 0$, such that $\|q_d\| \le d$, $\|\dot{q}_d\| \le q_1$, $\|\ddot{q}_d\| \le d_2$.

3. Main results

Firstly, we define the tracking error as

$$\begin{aligned} z_1 &= q_d - q \\ z_2 &= \alpha_1 - \dot{q} = \dot{z}_1 + K_1 z_1 \end{aligned} \tag{5}$$

where $\alpha_1 = K_1 z_1 + \dot{q}_d$ is an auxiliary variable.

Considering (3) and (5), the tracking error dynamics can be expressed as

$$\begin{aligned} \dot{z}_1 &= -K_1 z_1 + z_2 \\ \dot{z}_2 &= f\left(z^+\right) + g\left(z_1, q_d\right)\tau \end{aligned} \tag{6}$$

where $z = \left[z_1^T, z_2^T\right]^T$ represents the tracking error, $z^+ = \left[z_1^T, z_2^T, q_d^T, \dot{q}_d^T, \ddot{q}_d^T\right]^T$ represents the augmented error, $g(z_1, q_d) = -M^{-1}(q)$, $f\left(z^+\right) = M^{-1}(q)\left[C(q,\dot{q})\dot{q} + G(q)\right] + \dot{\alpha}_1$.

Define the sequence $\{t_k\}, k = 0, 1, \dots$.. as the solving time for MPC, in which $t_0 = 0$. The solving interval is expressed as $\Delta t = t_{k+1} - t_k$. Then the basic MPC strategy, for $s \in [t_k, t_k + T)$, is introduced as

$$\min_{\tau} J(z) = \int_{t_k}^{t_k+T} Q(z) + U(\tau)\,ds + \Psi\left(z\left(t_k + T\right)\right) \tag{7}$$

$$s.t.\begin{cases} z_j(t_k|t_k) = z_j(t_k), j = 1, 2 \\ \dot{z}_1(s) = -K_1 z_1 + z_2 \\ \dot{z}_2(s) = f\left(z^+\right) + g\left(z_1, q_d\right)\tau \\ |\tau_i| \le \lambda, i = 1, 2, \dots, n \end{cases} \tag{8}$$

where T is the prediction horizon, $Q(z)$ and $U(\tau)$ are positive-definite functions about tracking error z and input τ, respectively. The optimization problem (7) will be solved at time instant t_k under current initial stations $z_j(t_k|t_k), j = 1, 2$. The optimal or suboptimal torque $\tau^*(t)$ over $t \in [t_k, t_k + T)$ is obtained, and the first portion of $\tau^*(t)$ is implemented to the robotic system. Then the optimization problem over $t \in [t_{k+1}, t_{k+1} + T)$ is revisited at time instant t_{k+1} under new initial stations.

In practice, the basic MPC (7) (8) may not be realizable for robotic manipulators without specific design. Firstly, the accurate tracking error dynamics (6) might be unattainable since the uncertainties exist in $M^{-1}(q)$, $C(q,\dot{q})$ and $G(q)$. Then the solution of the nonlinear optimization problem and stability analysis are difficult for the unknown robotic system. To overcome above challenges, two groups of NNs are utilized under the proposed MPC structure: (i) the NN-based estimation model of the robotic system is established as a predictive model for approximating uncertain system dynamics online; (ii) actor–critic networks are established for solving the nonlinear optimization problem of MPC based on the predictive model. The stability of the closed-loop system is ensured at the same time.

3.1. NN-based predictive model

In this section, adaptive NNs are used as the predictive model to approximate the tracking error dynamics (6). Assume that τ is a persistently excited and feasible input which satisfies (4). The tracking error dynamics can be expressed as

$$\begin{aligned} \dot{z}_1 &= -K_1 z_1 + z_2 \\ \dot{z}_2 &= W_f^{*T}\varphi_f\left(z^+\right) + W_g^{*T}\varphi_g\left(z_1, q_d\right)\tau + \xi_m \end{aligned} \tag{9}$$

where $\xi_m = \xi_f + \xi_g\tau$ is the estimation error, $W_f^{*T}\varphi_f\left(z^+\right) + \xi_f = f\left(z^+\right)$, $W_g^{*T}\varphi_g\left(z_1, q_d\right) + \xi_g = g\left(z_1, q_d\right)$, $\varphi_f\left(z^+\right)$, $\varphi_g\left(z_1, q_d\right)$ are NN's activation functions which are selected as Gaussian in this paper.

Assumption 2 ([2,43]). The optimal NN's weights W_f^*, W_g^*, activation functions $\varphi_f\left(z^+\right)$, $\varphi_g\left(z_1, q_d\right)$, and approximation errors ξ_f, ξ_g are bounded, i.e. there exist $w_{f0} > 0$, $w_{g0} > 0$, $\varphi_{f0} > 0$, $\varphi_{g0} > 0$, $\xi_{f0} > 0$, $\xi_{g0} > 0$, such that $\sum_{i=1}^n \left\|W_{f,i}^*\right\| \le w_{f0}$, $\sum_{i=1}^n \left\|W_{g,i}^*\right\| \le w_{g0}$, $\left\|\varphi_f\left(z^+\right)\right\| \le \varphi_{f0}$, $\left\|\varphi_g\left(z_1, q_d\right)\right\| \le \varphi_{g0}$, $\left\|\xi_f\right\| \le \xi_{f0}$, $\left\|\xi_g\right\| \le \xi_{g0}$.

Lemma 5. *Let Assumption 2 and constraints (4) hold, then the estimation error ξ_m is bounded, i.e. there exists $\xi_{m0} > 0$, such that $\|\xi_m\| \le \xi_{m0}$.*

In the MPC scheme, the optimization problem is solved at the time sequence $\{t_k\}$ discretely. Thus the predictive model over $t \in [t_k, t_{k+1})$ can be defined as

$$\begin{aligned} \dot{\bar{z}}_1 &= -K_1\bar{z}_1 + \bar{z}_2 + L\Delta\bar{z}_1^k \\ \dot{\bar{z}}_2 &= \bar{f}\left(\bar{z}^+\right) + \bar{g}\left(\bar{z}_1, q_d\right)\tau + L\Delta\bar{z}_2^k \end{aligned} \tag{10}$$

where $\bar{g}\left(\bar{z}_1, q_d\right) = \hat{W}_{gk}^T\varphi_g\left(\bar{z}_1, q_d\right)$, $\bar{f}\left(\bar{z}^+\right) = \hat{W}_{fk}^T\varphi_f\left(\bar{z}^+\right)$, $\hat{W}_{fk}$ and $\hat{W}_{gk}$ are approximations of W_f^* and W_g^* at time t_k, respectively. L is a positive constant, $\Delta\bar{z}_j^k = z_j(t_k) - \bar{z}_j\left(t_k^-\right), j = 1, 2$ is the estimation error of z_j at time t_k. Define $\Delta\bar{z}_j^0 \equiv 0$.

NN's weights $\hat{W}_{fk}$ and $\hat{W}_{gk}$ in (10) are constants over $t \in [t_k, t_{k+1}), k = 0, 1, 2, \dots$., and are updated at time instant t_{k+1} through:

$$\begin{aligned} \hat{W}_{f(k+1),i} &= \hat{W}_{fk,i} + \Delta\hat{W}_{fk,i} \\ &= \hat{W}_{fk,i} + \alpha_f\left[-(1+\Delta t)\,\Theta_{fk}L\Delta t\Delta\bar{z}_{2,i}^k - k_f\hat{W}_{fk,i}\right] \end{aligned} \tag{11}$$

$$\begin{aligned} \hat{W}_{g(k+1),i} &= \hat{W}_{gk,i} + \Delta\hat{W}_{gk,i} \\ &= \hat{W}_{gk,i} + \alpha_g\left[-(1+\Delta t)\,\Theta_{gk}L\Delta t\Delta\bar{z}_{2,i}^k - k_g\hat{W}_{gk,i}\right] \end{aligned} \tag{12}$$

where $\alpha_f > 0$, $\alpha_g > 0$ are learning rates, $k_f > 0$, $k_g > 0$ are introduced for improving the robustness, $\Theta_{gk} \equiv \frac{1}{2}\sum_{l=1}^{N-1}\left(\varphi_g\left(\bar{z}_{1l}, q_{dl}\right)\tau_l + \varphi_g\left(\bar{z}_{1(l+1)}, q_{d(l+1)}\right)\tau_{l+1}\right)\Delta t_l$, $\Theta_{fk} \equiv \frac{1}{2}\sum_{l=1}^{N-1}\left(\varphi_f\left(\bar{z}_l^+\right) + \varphi_f\left(\bar{z}_{l+1}^+\right)\right)\Delta t_l$, in which $(.)_l = (.)(t_l)$, $\Delta t_l = t_{l+1} - t_l$, N is a finite positive integer, $t_1 = t_k$, and $t_N = t_{k+1}$.

Lemma 6. *Let Assumption 2 and constraints (4) hold, input τ satisfies PE condition. Then signals Θ_{fk}, Θ_{gk} are bounded, i.e. there exist $\theta_{f0} > 0$, $\theta_{g0} > 0$, such that $\left\|\Theta_{fk}\right\| \le \theta_{f0}$, $\left\|\Theta_{gk}\right\| \le \theta_{g0}$. Θ_{fk} and Θ_{gk} are persistently existed, too.*

According to the characteristic of MPC, the predictive tracking error $\bar{z}_j$ is updated by real value z_j at time instant t_{k+1} through

$$\bar{z}_j\left(t_{k+1}^+\right) = z_j\left(t_{k+1}\right), j = 1, 2 \tag{13}$$

Further more, it can be gotten that $\forall t \in [t_k, t_{k+1})$, it holds that $\Delta\bar{z}_2(t)^T\Delta\bar{z}_2(t) \leqslant \Delta\bar{z}_2^{(k+1)T}\Delta\bar{z}_2^{(k+1)}$.

Define $\tilde{W}_{fk} = W_f^* - \hat{W}_{fk}$, $\tilde{W}_{gk} = W_f^* - \hat{W}_{gk}$ as NN's weights estimation errors. For $t \in [t_k, t_{k+1})$, we consider $\Delta\bar{z}^k(t) = \Delta\bar{z}^k$,

$\tilde{W}_{fk,i}(t) = \tilde{W}_{fk,i}$ and $\tilde{W}_{gk,i}(t) = \tilde{W}_{gk,i}$. In the following theorem, the designed adaptive strategies (11) and (12) are shown to guarantee the boundedness of the predictive estimation error and NN's weights estimation errors.

Theorem 1. *Let the predictive model be defined as (10) and input signals satisfy constraints (4) and PE condition. Assume that Assumption 2 and Lemmas 1, 2, 4, 5, 6 hold. NN's weights $\hat{W}_{fk}$ and $\hat{W}_{gk}$ are updated through (11) and (12) at time instant $t_{k+1}, k = 0, 1, 2, \ldots$. The predictive tracking error $\bar{z}_j$ is updated by real value through (13) at time instant t_{k+1}. Then the predictive estimation error $\Delta\bar{z}^k$ and NN's weights estimation errors $\tilde{W}_{fk}$, $\tilde{W}_{gk}$ remain UUB, if it holds that:*

(i) $K_1 - 1 > 0$

(ii) $1 - L^2\Delta t > 0$

(iii) $1 - 2(1+\Delta t)L^2\Delta t^2 - 3(1+\Delta t)^2L^2\Delta t^2\left(\alpha_f\theta_{f0}^2 + \alpha_g\theta_{g0}^2\right) > 0$

(iv) $k_f - 3\alpha_f k_f^2 - 3(1+\Delta t)\theta_{f0}^2 > 0$

(v) $k_g - 3\alpha_g k_g^2 - 3(1+\Delta t)\theta_{g0}^2 > 0$

Proof. Construct a Lyapunov function candidate as

$$\begin{aligned}V_{mk} =& \frac{1}{2}\Delta\bar{z}_1^{kT}\Delta\bar{z}_1^k + \frac{1}{2}\Delta\bar{z}_2^{kT}\Delta\bar{z}_2^k \\ &+ \sum_{i=1}^n \frac{1}{2\alpha_f}\tilde{W}_{fk,i}^T\tilde{W}_{fk,i} + \sum_{i=1}^n \frac{1}{2\alpha_g}\tilde{W}_{gk,i}^T\tilde{W}_{gk,i} \\ =& V_{mz} + V_{mf} + V_{mg}\end{aligned} \tag{14}$$

where $V_{mz} = \frac{1}{2}\Delta\bar{z}_1^{kT}\Delta\bar{z}_1^k + \frac{1}{2}\Delta\bar{z}_2^{kT}\Delta\bar{z}_2^k$, $V_{mf} = \sum_{i=1}^n \frac{1}{2\alpha_f}\tilde{W}_{fk,i}^T\tilde{W}_{fk,i}$, $V_{mg} = \sum_{i=1}^n \frac{1}{2\alpha_g}\tilde{W}_{gk,i}^T\tilde{W}_{gk,i}$.

For $t \in [t_k, t_{k+1})$, V_{mk} is a constant obviously. At time instant t_{k+1}, the difference of V_{mk} can be calculated through

$$\Delta V_{mk} = \Delta V_{mz} + \Delta V_{mf} + \Delta V_{mg} \tag{15}$$

The first term in (15) is given by

$$\begin{aligned}\Delta V_{mz} =& \frac{1}{2}\Delta\bar{z}_1^{(k+1)T}\Delta\bar{z}_1^{(k+1)} - \frac{1}{2}\Delta\bar{z}_1^{kT}\Delta\bar{z}_1^k \\ &+ \frac{1}{2}\Delta\bar{z}_2^{(k+1)T}\Delta\bar{z}_2^{(k+1)} - \frac{1}{2}\Delta\bar{z}_2^{kT}\Delta\bar{z}_2^k\end{aligned} \tag{16}$$

in which

$$\begin{aligned}&\frac{1}{2}\Delta\bar{z}_1^{(k+1)T}\Delta\bar{z}_1^{(k+1)} \\ =& \int_{t_k}^{t_{k+1}} \Delta\bar{z}_1^T\Delta\dot{\bar{z}}_1 ds \le -(K_1-1)\int_{t_k}^{t_{k+1}} \Delta\bar{z}_1^T\Delta\bar{z}_1 ds \\ &+ \frac{1}{2}\int_{t_k}^{t_{k+1}} \Delta\bar{z}_2^T\Delta\bar{z}_2 ds + \frac{L^2}{2}\int_{t_k}^{t_{k+1}} \Delta\bar{z}_1^{kT}\Delta\bar{z}_1^k ds \\ \le& \frac{\Delta t}{2}\Delta\bar{z}_2^{(k+1)T}\Delta\bar{z}_2^{(k+1)} + \frac{\Delta t L^2}{2}\Delta\bar{z}_1^{kT}\Delta\bar{z}_1^k\end{aligned} \tag{17}$$

$$\begin{aligned}\Delta\bar{z}_2^{(k+1)} =& z_2(t_{k+1}) - \bar{z}_2\left(t_{k+1}^-\right) \\ =& \int_{t_k}^{t_{k+1}} \left(W_f^{*T}\varphi_f\left(z^+\right) - \hat{W}_{fk}^T\varphi_f\left(\bar{z}^+\right)\right) ds \\ &+ \int_{t_k}^{t_{k+1}} \left(W_g^{*T}\varphi_g\left(z_1, q_d\right)\tau - \hat{W}_{gk}^T\varphi_g\left(\bar{z}_1, q_d\right)\tau\right) ds \\ &+ \int_{t_k}^{t_{k+1}} \xi_m ds - \int_{t_k}^{t_{k+1}} L\Delta\bar{z}_2^k ds\end{aligned} \tag{18}$$

From Lemmas 1 and 2, we have

$$\begin{aligned}&\int_{t_k}^{t_{k+1}} \left(W_f^{*T}\varphi_f\left(z^+\right) - \hat{W}_{fk}^T\varphi_f\left(\bar{z}^+\right)\right) ds \\ =& \int_{t_k}^{t_{k+1}} \left(W_f^{*T}\left(\varphi_f\left(z^+\right) - \varphi_f\left(\bar{z}^+\right)\right) + \tilde{W}_{fk}^T\varphi_f\left(\bar{z}^+\right)\right) ds \\ =& \varpi_f + \tilde{W}_{fk}^T\Theta_{fk}\end{aligned} \tag{19}$$

where $\varpi_f = W_f^{*T}\left(\varphi_f\left(z^+\left(\varepsilon_f\right)\right) - \varphi_f\left(\bar{z}^+\left(\varepsilon_f\right)\right)\right)\Delta t + \zeta_f$, ζ_f is the integral error, $\varepsilon_f \in [t_k, t_{k+1})$, Θ_{fk} is defined earlier. Analogously,

$$\begin{aligned}&\int_{t_k}^{t_{k+1}} \left(W_g^{*T}\varphi_g\left(z_1, q_d\right)\tau - \hat{W}_{gk}^T\varphi_g\left(\bar{z}_1, q_d\right)\tau\right) ds \\ =& \varpi_g + \tilde{W}_{gk}^T\Theta_{gk}\end{aligned} \tag{20}$$

where $\varpi_g = W_g^{*T}\left(\varphi_g\left(z_1\left(\varepsilon_g\right), q_d\left(\varepsilon_g\right)\right) - \varphi_g\left(\bar{z}_1\left(\varepsilon_g\right), q_d\left(\varepsilon_g\right)\right)\right)\tau\left(\varepsilon_g\right)\Delta t + \zeta_g$, ζ_g is the integral error, $\varepsilon_g \in [t_k, t_{k+1})$, Θ_{gk} is defined earlier.

$$\begin{aligned}&\int_{t_k}^{t_{k+1}} \xi_m - L\Delta\bar{z}_2^k ds \\ =& \xi_m\left(\varepsilon_\xi\right)\Delta t - L\Delta t\Delta\bar{z}_2^k = \varpi_\xi - L\Delta t\Delta\bar{z}_2^k\end{aligned} \tag{21}$$

where $\varepsilon_\xi \in [t_k, t_{k+1})$.

Substituting (19)–(21) into (18), we have

$$\Delta\bar{z}_2^{k+1} = \varpi + \tilde{W}_{fk}^T\Theta_{fk} + \tilde{W}_{gk}^T\Theta_{gk} - L\Delta t\Delta\bar{z}_2^k \tag{22}$$

in which $\varpi = \varpi_f + \varpi_g + \varpi_\xi$.

Lemma 7. *Let Lemmas 2, 5, 6, Assumption 2 and constraints (4) hold, then the signal ϖ is bounded, i.e. there exists $\varpi_0 > 0$, such that $\|\varpi\| \le \varpi_0$.*

Substituting (17) (22) into (16), it can be gotten

$$\begin{aligned}\Delta V_{mz} \le& -\frac{1}{2}\left(1 - 2(1+\Delta t)L^2\Delta t^2\right)\Delta\bar{z}_2^{kT}\Delta\bar{z}_2^k \\ &- \frac{1}{2}\left(1 - L^2\Delta t\right)\Delta\bar{z}_1^{kT}\Delta\bar{z}_1^k + 2(1+\Delta t)\|\varpi_0\|^2 \\ &- (1+\Delta t)L\Delta t\Delta\bar{z}_2^{kT}\left(\tilde{W}_{fk}^T\Theta_{fk} + \tilde{W}_{gk}^T\Theta_{gk}\right) \\ &+ \frac{3(1+\Delta t)}{2}\theta_{f0}^2\sum_{i=1}^n\left\|\tilde{W}_{fk,i}\right\|^2 \\ &+ \frac{3(1+\Delta t)}{2}\theta_{g0}^2\sum_{i=1}^n\left\|\tilde{W}_{gk,i}\right\|^2\end{aligned} \tag{23}$$

The second term in (15) is given by

$$\begin{aligned}\Delta V_{mf} =& \sum_{i=1}^n \frac{1}{2\alpha_f}\tilde{W}_{f(k+1),i}^T\tilde{W}_{f(k+1),i} - \frac{1}{2\alpha_f}\tilde{W}_{fk,i}^T\tilde{W}_{fk,i} \\ =& \sum_{i=1}^n \frac{1}{2\alpha_f}\Delta\hat{W}_{fk,i}^T\Delta\hat{W}_{fk,i} - \frac{1}{\alpha_f}\tilde{W}_{fk,i}^T\Delta\hat{W}_{fk,i} \\ \le& \frac{1}{2}\sum_{i=1}^n\left(k_f + 3\alpha_f k_f^2\right)\left\|W_{f,i}^*\right\|^2 \\ &- \frac{1}{2}\sum_{i=1}^n\left(k_f - 3\alpha_f k_f^2\right)\left\|\tilde{W}_{fk,i}\right\|^2 \\ &+ \frac{3\alpha_f}{2}(1+\Delta t)^2L^2\Delta t^2\theta_{f0}^2\left\|\Delta\bar{z}_2^k\right\|^2 \\ &+ (1+\Delta t)\sum_{i=1}^n\tilde{W}_{fk,i}^T\Theta_{fk}L\Delta t\Delta\bar{z}_{2,i}^k\end{aligned} \tag{24}$$

Analogously, the third term in (15) is expressed as

$$\begin{aligned}\Delta V_{mg} =& \sum_{i=1}^{n} \frac{1}{2\alpha_g} \tilde{W}_{g(k+1),i}^T \tilde{W}_{g(k+1),i} - \frac{1}{2\alpha_g} \tilde{W}_{gk,i}^T \tilde{W}_{gk,i} \\ \leq& \frac{1}{2} \sum_{i=1}^{n} \left(k_g + 3\alpha_g k_g^2\right) \left\|W_{g,i}^*\right\|^2 \\ &- \frac{1}{2} \sum_{i=1}^{n} \left(k_g - 3\alpha_g k_g^2\right) \left\|\tilde{W}_{gk,i}\right\|^2 \\ &+ \frac{3\alpha_g}{2} (1+\Delta t)^2 L^2 \Delta t^2 \theta_{g0}^2 \left\|\Delta \bar{z}_2^k\right\|^2 \\ &+ (1+\Delta t) \sum_{i=1}^{n} \tilde{W}_{gk,i}^T \Theta_{gk} L \Delta t \Delta \bar{z}_{2,i}^k \end{aligned} \tag{25}$$

Substituting (23)–(25) into (15), and noting that

$$L\Delta t \Delta \bar{z}_{2k}^T \tilde{W}_{fk}^T \Theta_{fk} = \sum_{i=1}^{n} \tilde{W}_{fk,i}^T \Theta_{fk} L \Delta t \Delta \bar{z}_{2k,i} \tag{26}$$

$$L\Delta t \Delta \bar{z}_{2k}^T \tilde{W}_{gk}^T \Theta_{gk} = \sum_{i=1}^{n} \tilde{W}_{gk,i}^T \Theta_{gk} L \Delta t \Delta \bar{z}_{2k,i} \tag{27}$$

We have

$$\begin{aligned}\Delta V_{mk} \leq& -\frac{1}{2}\kappa_1 \Delta \bar{z}_1^{kT} \Delta \bar{z}_1^k - \frac{1}{2}\kappa_2 \Delta \bar{z}_2^{kT} \Delta \bar{z}_2^k \\ &- \sum_{i=1}^{n} \frac{\kappa_3}{2} \left\|\tilde{W}_{fk,i}\right\|^2 - \sum_{i=1}^{n} \frac{\kappa_4}{2} \left\|\tilde{W}_{gk,i}\right\|^2 + C_m \\ \leq& -\kappa V_{mk} + C_m \end{aligned} \tag{28}$$

where $\kappa_1 = 1 - L^2\Delta t$, $\kappa_2 = 1 - 2(1+\Delta t) L^2 \Delta t^2 - 3(1+\Delta t)^2 L^2 \Delta t^2 \left(\alpha_f \theta_{f0}^2 + \alpha_g \theta_{g0}^2\right)$, $\kappa_3 = k_f - 3\alpha_f k_f^2 - 3(1+\Delta t)\theta_{f0}^2$, $\kappa_4 = k_g - 3\alpha_g k_g^2 - 3(1+\Delta t)\theta_{g0}^2$, $\kappa = \min\left(\kappa_1, \kappa_2, \alpha_f\kappa_3, \alpha_g\kappa_4\right)$, $C_m = 2(1+\Delta t)\|\varpi_0\|^2 + \frac{1}{2}\sum_{i=1}^{n}\left(k_f + 3\alpha_f k_f^2\right)\left\|W_{f,i}^*\right\|^2 + \frac{1}{2}\sum_{i=1}^{n}\left(k_g + 3\alpha_g k_g^2\right)\left\|W_{g,i}^*\right\|^2$.

The parameters fulfill that $\kappa_1 > 0$, $\kappa_2 > 0$, $\kappa_3 > 0$ and $\kappa_4 > 0$. According to Lemma 4, the predictive estimation error $\Delta \bar{z}_j^k$ and NN's weights errors $\tilde{W}_{fk}$, $\tilde{W}_{gk}$ will remain UUB.

It is worth noting that the PE condition of Θ_{fk} and Θ_{gk} ensures sufficient signals of the estimation error space to keep $\hat{W}_{fk}$ and $\hat{W}_{gk}$ from converging to zero.

Furthermore, the estimation error $\Delta \bar{z}^k$ converges asymptotically to the compact set $\Omega_{\Delta\bar{z}} := \left\{\Delta \bar{z}^k \in \Re^{2n} \middle| \left\|\Delta \bar{z}^k\right\| \leqslant \sqrt{\mathbb{Z}}\right\}$, where $\mathbb{Z} = 2\left(V_{mk}^*(0) + \frac{C_m}{\kappa}\right)$. The proof can refer to [41].

Based on the predictive model (10) with adaptive laws (11) and (12), the MPC strategy (7) (8), for $s \in [t_k, t_k+T]$, is reformulated as

$$\min_{\bar{\tau}} J(\bar{z}) = \int_{t_k}^{t_k+T} Q(\bar{z}) + U(\bar{\tau})\, ds + \Psi(\bar{z}(t_k+T)) \tag{29}$$

$$s.t. \begin{cases} \bar{z}_j(t_k|t_k) = z_j(t_k), j = 1, 2 \\ \dot{\bar{z}}_1(s) = -K_1\bar{z}_1 + \bar{z}_2 + L\Delta \bar{z}_1^k \\ \dot{\bar{z}}_2(s) = \bar{f}(\bar{z}^+) + \bar{g}(\bar{z}_1, q_d)\tau + L\Delta \bar{z}_2^k \\ |\bar{\tau}_i| \leq \lambda, i = 1, 2, \ldots, n \end{cases} \tag{30}$$

where parameters have the same definitions with (7) (8).

3.2. Optimization problem solving for MPC

In this section, we utilize optimal control strategy and Hamiltonian function to solve (29) subject to (30) for $s \in [t_k, t_k+T)$. Both optimal tracking performance and stability analysis of the predictive system are considered.

For $t \in [t_k, t_k+T)$, the cost function is redefined as

$$\bar{J}_t(\bar{z}) = \int_{t}^{t_k+T} Q(\bar{z}) + U(\bar{\tau})\, ds + \Psi(\bar{z}(t_k+T)) \tag{31}$$

where $Q(\bar{z}) = \bar{z}_1^T Q_1 \bar{z}_1 + \bar{z}_2^T Q_2 \bar{z}_2$, Q_1 and $Q_2 \in \Re^{n\times n}$ are symmetric positive definite. $U(\bar{\tau}) = \int_0^{\bar{\tau}} \lambda \beta^{-1}(v/\lambda)^T R dv$, $\beta(.) = \tanh(.)$, $R = diag(r_1, \ldots, r_n) > 0$. $U(\bar{\tau})$ is defined as an integrand function for ensuring that the input constraints (4) can be satisfied. $\Psi(\bar{z}(t_k+T))$ is the terminal penalty which can be seen as an estimation of the optimal cost function from time instant t_k+T to infinity.

Assume that the cost function is smooth. Then the optimal cost function can be expressed as

$$\bar{J}_t^* = W^{*T}\varphi_c(\bar{z}) + \xi_c \tag{32}$$

where ξ_c is the estimation error, $\varphi_c(\bar{z})$ is the NN's activation function which is selected as Gaussian in this paper. The terminal penalty can be defined as $\Psi(\bar{z}(t_k+T)) = W^{*T}\varphi_c(\bar{z}(t_k+T)) + \xi_c$. Then the optimal cost function's gradient can be expressed as

$$\begin{aligned}\nabla_j \bar{J}_t^* =& \frac{\partial \bar{J}_t^*(\bar{z})}{\partial \bar{z}_j} = \left(\frac{\partial \varphi_c(\bar{z})}{\partial \bar{z}_j}\right)^T W^* + \frac{\partial \xi_c}{\partial \bar{z}_j} \\ =& \nabla_j \varphi_c(\bar{z})^T W^* + \nabla_j \xi_c, j = 1, 2\end{aligned} \tag{33}$$

Assumption 3 ([43]). The optimal NN's weights W^*, activation function $\varphi_c(\bar{z})$ and its gradients $\nabla_1\varphi_c(\bar{z})$, $\nabla_2\varphi_c(\bar{z})$, estimation error ξ_c and its gradients $\nabla_1\xi_c$, $\nabla_2\xi_c$ are bounded, i.e. there exist $w_{c0} > 0$, $\varphi_{c0} > 0$, $\varphi_{d1c0} > 0$, $\varphi_{d2c0} > 0$, $\xi_{c0} > 0$, $\xi_{d1c0} > 0$, $\xi_{d2c0} > 0$, such that $\|W^*\| \leq w_{c0}$, $\|\varphi_c(\bar{z})\| \leq \varphi_{c0}$, $\|\nabla_1\varphi_c(\bar{z})\| \leq \varphi_{d1c0}$, $\|\nabla_2\varphi_c(\bar{z})\| \leq \varphi_{d2c0}$, $\|\xi_c\| \leq \xi_{c0}$, $\|\nabla_1\xi_c\| \leq \xi_{d1c0}$, $\|\nabla_2\xi_c\| \leq \xi_{d2c0}$.

According to the optimal control theory, the Hamiltonian function can be expressed as

$$\bar{H}\left(\bar{z}, \bar{\tau}(\bar{z}), \nabla \bar{J}_t^*\right) = \nabla_1 \bar{J}_t^{*T}\dot{\bar{z}}_1 + \nabla_2 \bar{J}_t^{*T}\dot{\bar{z}}_2 + Q(\bar{z}) + U(\bar{\tau}) \tag{34}$$

Under the stationarity condition, the optimal control strategy over $t \in [t_k, t_k+T)$ can be calculated by

$$\bar{\tau}^*(t) = \arg\min_{\bar{\tau}} \bar{H}\left(\bar{z}, \bar{\tau}(\bar{z}), \nabla \bar{J}_t^*\right) = -\lambda \tanh\left(\bar{\mathrm{T}}^*\right) \tag{35}$$

where $\bar{\mathrm{T}}^* = \frac{1}{\lambda}R^{-1}\bar{g}^T(\bar{z}_1, q_d)\nabla_2 \bar{J}_t^* = \frac{1}{\lambda}R^{-1}\bar{g}^T(\bar{z}_1, q_d)\left(\nabla_2\varphi_c(\bar{z})^T W^* + \nabla_2\xi_c\right)$.

Substituting (35) into $U(\bar{\tau})$, we have

$$\begin{aligned}U\left(\bar{\tau}^*\right) =& \lambda\left(\tanh^{-1}\left(\bar{\tau}^*/\lambda\right)\right)^T R\bar{\tau}^* \\ &+ \frac{1}{2}\lambda^2 \breve{R} \ln\left(\bar{1} - \Xi^2\left(\tanh\left(\bar{\mathrm{T}}^*\right)\right)\right) \\ =& \lambda \nabla_2 \bar{J}_t^{*T} \bar{g}(\bar{z}_1, q_d) \tanh\left(\bar{\mathrm{T}}^*\right) \\ &- \lambda^2 \breve{R} \ln\left(\cosh\left(\bar{\mathrm{T}}^*\right)\right)\end{aligned} \tag{36}$$

where $\breve{R} = [r_1, \ldots, r_n] \in \Re^{1\times n}$, $\bar{1} = [1, \ldots, 1]^T \in \Re^{n\times 1}$, $\Xi^2(.)$ is defined as an operation which squares each element of (.) severally.

Substituting (35) (36) into (34), the optimal Hamiltonian function can be expressed as

$$\begin{aligned}&\bar{H}^*\left(\bar{z}, \bar{\tau}^*, \nabla \bar{J}_t^*\right) \\ =& \nabla_1 \bar{J}_t^{*T}\dot{\bar{z}}_1 + \nabla_2 \bar{J}_t^{*T}\dot{\bar{z}}_2 + Q(\bar{z}) + U\left(\bar{\tau}^*\right) \\ =& W^{*\mathrm{T}}\nabla\varphi_c(\bar{z})\bar{F} + Q(\bar{z}) - \lambda^2 \breve{R}\ln\left(\cosh\left(\bar{\mathrm{T}}^*\right)\right) + \varepsilon_{HJB} \\ =& 0\end{aligned} \tag{37}$$

where $\bar{F} = \begin{bmatrix} -K_1\bar{z}_1 + \bar{z}_2 + L\Delta\bar{z}_1^k \\ \bar{f}\left(\bar{z}^+\right) + L\Delta\bar{z}_2^k \end{bmatrix}$, $\nabla\varphi_c(\bar{z}) = \left[\nabla_1\varphi_c(\bar{z})\ \ \nabla_2\varphi_c(\bar{z})\right]$, $\varepsilon_{HJB} = \nabla_2\xi_c{}^T\left(\bar{f}\left(\bar{z}^+\right) + L\Delta\bar{z}_2^k\right) + \nabla_1\xi_c{}^T\left(-K_1\bar{z}_1 + \bar{z}_2 + L\Delta\bar{z}_1^k\right)$.

For remaining optimal tracking performance and stability of the predictive system, as well as taking advantage of prior knowledge of the predictive model, we design actor–critic networks, which have the same activation function but different weights, to approximate the control strategy (35) and cost function (32). The critic network is defined as

$$\hat{\bar{J}}_t = \hat{W}_c^T \varphi_c\left(\bar{z}\right) \tag{38}$$

with the terminal penalty is expressed as $\hat{\Psi}\left(\bar{z}\left(t_k+T\right)\right) = \hat{W}_c^T \varphi_c\left(\bar{z}\left(t_k+T\right)\right)$. $\hat{W}_c$ is the estimation of W^* in the cost function.

The actor network is defined as $\hat{N}_a = \hat{W}_a^T \varphi_c\left(\bar{z}\right)$. The estimation control strategy can be expressed as

$$\hat{\bar{\tau}}\left(t\right) = \arg\min_{\bar{\tau}} \bar{H}\left(\bar{z}, \bar{\tau}\left(\bar{z}\right), \nabla \hat{N}_a\right) = -\lambda \tanh\left(\hat{\bar{\mathrm{T}}}\right) \tag{39}$$

where $\hat{\bar{\mathrm{T}}} = \frac{1}{\lambda} R^{-1} \bar{g}^T\left(\bar{z}_1, q_d\right) \nabla_2 \varphi_c\left(\bar{z}\right)^T \hat{W}_a$. $\hat{W}_a$ is the estimation of W^* in the control strategy. The estimation errors of actor–critic networks can be defined as $\tilde{W}_a = W^* - \hat{W}_a$, $\tilde{W}_c = W^* - \hat{W}_c$.

Substituting (38) (39) into (34), the estimation of Hamiltonian function can be gotten:

$$\begin{aligned}
&\hat{\bar{H}}\left(\bar{z}, \hat{\bar{\tau}}, \nabla \hat{\bar{J}}_t\right) \\
=&\nabla_1 \hat{\bar{J}}_t^T \dot{\bar{z}}_1 + \nabla_2 \hat{\bar{J}}_t^T \dot{\bar{z}}_2 + Q\left(\bar{z}\right) + U\left(\hat{\bar{\tau}}\right) \\
=&\hat{W}_c^T \left(\nabla \varphi_c\left(\bar{z}\right) \bar{F} + \nabla_2 \varphi_c\left(\bar{z}\right) \bar{g}\left(\bar{z}_1, q_d\right) \hat{\bar{\tau}}\right) + Q\left(\bar{z}\right) \\
&- \hat{W}_a^T \nabla_2 \varphi_c\left(\bar{z}\right) \bar{g}\left(\bar{z}_1, q_d\right) \hat{\bar{\tau}} - \lambda^2 \breve{R} \ln\left(\cosh\left(\hat{\bar{\mathrm{T}}}\right)\right)
\end{aligned} \tag{40}$$

Introducing (37), the estimation error of $\bar{H}$ can be defined as

$$\begin{aligned}
e_{\bar{H}} =&\hat{\bar{H}}\left(\bar{z}, \hat{\bar{\tau}}, \nabla \hat{\bar{J}}_t\right) \\
=&\hat{\bar{H}}\left(\bar{z}, \hat{\bar{\tau}}, \nabla \hat{\bar{J}}_t\right) - \bar{H}^*\left(\bar{z}, \bar{\tau}^*, \nabla \bar{J}_t^*\right) \\
=&-\tilde{W}_c^T w + \tilde{W}_a^T \nabla_2 \varphi_c\left(\bar{z}\right) \bar{g}\left(\bar{z}_1, q_d\right) \hat{\bar{\tau}} - \varepsilon_{HJB} \\
&+ \lambda^2 \breve{R}\left(\ln\left(\cosh\left(\bar{\mathrm{T}}^*\right)\right) - \ln\left(\cosh\left(\hat{\bar{\mathrm{T}}}\right)\right)\right)
\end{aligned} \tag{41}$$

where $w = \nabla \varphi_c\left(\bar{z}\right) \bar{F} + \nabla_2 \varphi_c\left(\bar{z}\right) \bar{g}\left(\bar{z}_1, q_d\right) \hat{\bar{\tau}}$.

The nonlinear term $\breve{R}\left(\ln\left(\cosh\left(\bar{\mathrm{T}}^*\right)\right) - \ln\left(\cosh\left(\hat{\bar{\mathrm{T}}}\right)\right)\right)$ in (41) can be expanded into a Taylor series as

$$\begin{aligned}
&\lambda^2 \breve{R}\left(\ln\left(\cosh\left(\bar{\mathrm{T}}^*\right)\right) - \ln\left(\cosh\left(\hat{\bar{\mathrm{T}}}\right)\right)\right) \\
=&\lambda^2 \breve{R} \frac{\partial \ln\left(\cosh\left(\hat{\bar{\mathrm{T}}}\right)\right)}{\partial \hat{\mathrm{T}}}\left(\bar{\mathrm{T}}^* - \hat{\bar{\mathrm{T}}}\right) + o\left(\hat{\bar{\mathrm{T}}}\right) \\
=&A \nabla_2 \varphi_c\left(\bar{z}\right)^T \tilde{W}_a + \varepsilon_o
\end{aligned} \tag{42}$$

where $A = \lambda \tanh^T\left(\hat{\bar{\mathrm{T}}}\right) \bar{g}^T\left(\bar{z}_1, q_d\right)$, $\varepsilon_o = A \nabla_2 \xi_c + o\left(\hat{\bar{\mathrm{T}}}\right)$, $o\left(\hat{\bar{\mathrm{T}}}\right)$ is the high order term. Obviously, the vector A is bounded by $A_{max} > 0$, which means $\|A\| \leq A_{\max}$.

Assumption 4. The high order term $o\left(\hat{\bar{\mathrm{T}}}\right)$ is bounded, i.e. there exists $o_0 > 0$, such that $\left\|o\left(\hat{\bar{\mathrm{T}}}\right)\right\| \leq o_0$.

Lemma 8. *Let Assumptions 3, 4 hold. Then the error ε_o is bounded, i.e. there exists $\varepsilon_{o0} > 0$, such that $\|\varepsilon_o\| \leq \varepsilon_{o0}$.*

Substituting (42) into (41), we have

$$e_{\bar{H}} = -\tilde{W}_c^T w + \varepsilon_o - \varepsilon_{HJB} \tag{43}$$

For driving the result to converge to the optimal or suboptimal solution, and guaranteeing the stability of the predictive system, the updating law for $\hat{W}_c$ is developed as:

$$\begin{aligned}
\dot{\hat{W}}_c &= -\frac{\alpha_c}{\left(1 + w^T w\right)^2} \frac{\partial E_{\bar{H}}}{\partial \hat{W}_c} - 2\alpha_c k_s \hat{W}_c - \alpha_c k_p \hat{W}_a \\
&= -\frac{\alpha_c w}{\left(1 + w^T w\right)^2} e_{\bar{H}} - 2\alpha_c k_s \hat{W}_c - \alpha_c k_p \hat{W}_a
\end{aligned} \tag{44}$$

where $E_{\bar{H}} = \frac{1}{2}\left\|e_{\bar{H}}\right\|^2$ is an integral squared error, $\alpha_c > 0$ is the learning rate, $k_s > 0$, $k_p > 0$ are learning parameters. The first term of (44) is used to drive the error $e_{\bar{H}}$ to zero, the other terms are used to guarantee the stability of the predictive system.

Lemma 9 ([44]). *The normalized signal $\frac{w}{\left(1+w^T w\right)}$ is bounded. i.e. there exists $w_{\max} \in (0, 1)$, such that $\left\|\frac{w}{\left(1+w^T w\right)}\right\| \leq w_{\max}$.*

The updating law for $\hat{W}_a$ is designed as

$$\dot{\hat{W}}_a = \alpha_a k_p \hat{W}_c - 2\alpha_a k_a \hat{W}_a + \alpha_a \Upsilon_{W_a} \bar{z}_2 \tag{45}$$

where $\alpha_a > 0$ is the learning rate, $k_a > 0$ is the learning parameter, $\Upsilon_{W_a} = \nabla_2 \varphi_c\left(\bar{z}\right) \bar{g}\left(\bar{z}_1, q_d\right) R^{-1} diag\left(\Xi^2\left(sech\left(\hat{\bar{\mathrm{T}}}\right)\right)\right) \bar{g}^T\left(\bar{z}_1, q_d\right)$.

For the predictive model (10) with cost function (31), based on the estimation cost function (38) and control strategy (39) with adaptive laws (44) and (45), the MPC solving algorithm over $s \in [t_k, t_k + T)$ can be summarized as Algorithm 1. The maximum number of iterations N_R is adopted to guarantee the finite computation time for Algorithm 1.

Algorithm 1 MPC solving algorithm

Input:
$\hat{W}_{ck-}$, $\hat{W}_{ak-}$: NN's weights which are gotten from last period;
$z\left(t_k\right)$: Initial value of the predictive model;
$\alpha_c, \alpha_a, k_c, k_p, k_a$: Learning rates and parameters of actor–critic networks;
Δ: Convergence thresholds in actor–critic networks;
N_R: Maximum number of iterations

```
Initialize Ŵ_ck, Ŵ_ak by Ŵ_ck−, Ŵ_ak−;
Initialize N_R = 1;
repeat
    Initialize z̄(t_k) = z(t_k);
    for t ∈ [t_k, t_k + T) do
        Compute the control input τ̄̂(t) via (39);
        Compute the estimation Hamiltonian function via (40);
        update Ŵ_c, Ŵ_a via (44), (45);
        update predictive tracking error z̄ via (10);
    end for
    n_R = n_R + 1;
until ‖Ŵ_c − Ŵ_ck−‖ + ‖Ŵ_a − Ŵ_ak−‖ ≤ Δ or n_R > N_R
```

Return:
$\hat{W}_c$, $\hat{W}_a$, $\hat{\bar{\tau}}$

For stability analysis, we have the following assumption:

Assumption 5 ([45]). The predictive system with cost function (31) is asymptotic stability under the optimal control strategy (35). Define $\bar{V}_1 = \frac{1}{2}\bar{z}^T \bar{z}$ as a Lyapunov function, then there exist positive definite values Π_1, Π_2 which satisfy

$$\begin{aligned}
\dot{\bar{V}}_1^* &= \bar{z}_1^T \dot{\bar{z}}_1 + \bar{z}_2^T\left(\bar{f}\left(\bar{z}^+\right) + L\Delta \bar{z}_{2k} + \bar{g}\left(\bar{z}_1, q_d\right) \bar{\tau}^*\right) \\
&\leq -\Pi_1 \|\bar{z}_1\|^2 - \Pi_2 \|\bar{z}_2\|^2 \leq 0
\end{aligned} \tag{46}$$

In the following theorem, the MPC solving algorithm is shown to guarantee the boundedness of the predictive tracking error $\bar{z}$ and NN's weights errors $\tilde{W}_c$, $\tilde{W}_a$ for $t \in [t_k, t_k + T)$, $\forall k \in \mathbb{N}$.

Theorem 2. *Let Assumptions 3–5, Lemma 8 hold. Under the MPC solving algorithm, the predictive tracking error $\bar{z}$ and NN's weights errors $\tilde{W}_c$, $\tilde{W}_a$ remain UUB for $t \in [t_k, t_k+T)$, $\forall k \in \mathbb{N}$, if it holds that:*

(i) $e^{-\delta\Delta t}\underline{\lambda}_{Q_1} + \Pi_1 - 2\xi_{d1c0}^2\|K_1\|^2 > 0$

(ii) $e^{-\delta\Delta t}\underline{\lambda}_{Q_2} + \Pi_2 - \frac{1}{2}\bar{\lambda}_{\mathrm{B}} - 2\xi_{d1c0}^2 - \frac{1}{2} > 0$

(iii) $\underline{\lambda}_{\mathrm{M}_w} + k_s - \frac{\varepsilon_{o0}}{2} - \frac{k_p}{2} - \frac{w_{\max}^2}{2} > 0$

(iv) $k_a - \frac{k_p}{2} - \bar{\lambda}_{\breve{B}} > 0$

(v) $\delta > 0$

Proof. Construct a Lyapunov function candidate as

$$\begin{aligned}\bar{V} &= \frac{1}{2}\bar{z}^T\bar{z} + e^{-\delta(t-t_k)}\bar{J}_t^* + \frac{1}{2\alpha_c}\tilde{W}_c^T\tilde{W}_c + \frac{1}{2\alpha_a}\tilde{W}_a^T\tilde{W}_a \\ &= \bar{V}_1 + \bar{V}_2 + \bar{V}_3 + \bar{V}_4\end{aligned} \tag{47}$$

The time derivative of $\bar{V}$ is

$$\dot{\bar{V}} = \dot{\bar{V}}_1 + \dot{\bar{V}}_2 + \dot{\bar{V}}_3 + \dot{\bar{V}}_4 \tag{48}$$

The first term of (48) is given by

$$\begin{aligned}\dot{\bar{V}}_1 =& \bar{z}_1^T\dot{\bar{z}}_1 + \bar{z}_2^T\dot{\bar{z}}_2 \\ =& \bar{z}_1^T\dot{\bar{z}}_1 + \bar{z}_2^T\left(\bar{f}\left(\bar{z}^+\right) + L\Delta\bar{z}_{2k} + \bar{g}\left(\bar{z}_1, q_d\right)\bar{\tau}^*\right) \\ &- \bar{z}_2^T\bar{g}\left(\bar{z}_1, q_d\right)\bar{\tau}^* + \bar{z}_2^T\bar{g}\left(\bar{z}_1, q_d\right)\hat{\bar{\tau}}\end{aligned} \tag{49}$$

Substituting (35) (39) into the last term of (49), and expand it into a Taylor series as

$$\begin{aligned}&- \bar{z}_2^T\bar{g}\left(\bar{z}_1, q_d\right)\bar{\tau}^* + \bar{z}_2^T\bar{g}\left(\bar{z}_1, q_d\right)\hat{\bar{\tau}} \\ =& \bar{z}_2^T\bar{g}\left(\bar{z}_1, q_d\right)\left(\lambda\tanh\left(\bar{\mathrm{T}}^*\right) - \lambda\tanh\left(\hat{\bar{\mathrm{T}}}\right)\right) \\ =& \bar{z}_2^T B\nabla_2\varphi_c(\bar{z})^T\tilde{W}_a + \bar{z}_2^T B\nabla_2\xi_c + \bar{z}_2^T\bar{g}\left(\bar{z}_1, q_d\right)o_1\left(\hat{\bar{\mathrm{T}}}\right)\end{aligned} \tag{50}$$

where $\mathrm{B} = \bar{g}\left(\bar{z}_1, q_d\right)diag\left(\Xi^2\left(sech\left(\hat{\bar{\mathrm{T}}}\right)\right)\right)R^{-1}\bar{g}^T\left(\bar{z}_1, q_d\right)$, $o_1\left(\hat{\bar{\mathrm{T}}}\right)$ is the high order term. According to Theorem 1 and Assumption 4, signals $\bar{g}\left(\bar{z}_1, q_d\right)$ and $o_1\left(\hat{\bar{\mathrm{T}}}\right)$ are bounded, i.e. there exist $\hat{w}_{g\max} > 0$, $o_{10} > 0$, such that $\sum_{i=1}^n\|\bar{g}_i\| \le \hat{w}_{g\max}\varphi_{g0}$, $\left\|o_1\left(\hat{\bar{\mathrm{T}}}\right)\right\| \le o_{10}$.

Substituting (46) (50) into (49), we have

$$\begin{aligned}\dot{\bar{V}}_1 \le& -\Pi_1\|\bar{z}_1\|^2 - \left(\Pi_2 - \frac{1}{2}\bar{\lambda}_{\mathrm{B}} - \frac{1}{2}\right)\|\bar{z}_2\|^2 \\ &+ \bar{z}_2^T\mathrm{B}\nabla_2\varphi_c(\bar{z})^T\tilde{W}_a + C_{\bar{z}}\end{aligned} \tag{51}$$

where $C_{\bar{z}} = \frac{1}{2}\bar{\lambda}_{\mathrm{B}}\xi_{d2c0}^2 + \frac{1}{2}\hat{w}_{g\max}^2\varphi_{g0}^2 o_{10}^2$.

The second term of (48) is given by

$$\begin{aligned}\dot{\bar{V}}_2 =& -\delta e^{-\delta(t-t_k)}\bar{J}_t^* + e^{-\delta(t-t_k)}\dot{\bar{J}}_t^* = -\delta e^{-\delta(t-t_k)}\bar{J}_t^* \\ &+ e^{-\delta(t-t_k)}\left(\nabla_1\bar{J}_t^{*T}\dot{\bar{z}}_1 + \nabla_2\bar{J}_t^{*T}\left(\bar{f}\left(\bar{z}^+\right) + L\Delta\bar{z}_{2k}\right)\right) \\ &+ e^{-\delta(t-t_k)}\nabla_2\bar{J}_t^{*T}\bar{g}\left(\bar{z}_1, q_d\right)\hat{\bar{\tau}}\end{aligned} \tag{52}$$

Substituting (37) into (52), we have

$$\begin{aligned}\dot{\bar{V}}_2 \le& -\delta e^{-\delta(t-t_k)}\bar{J}_t^* - e^{-\delta\Delta t}\bar{z}_1^T Q_1\bar{z}_1 \\ &- e^{-\delta\Delta t}\bar{z}_2^T Q_2\bar{z}_2 + \tilde{W}_a^T\breve{B}\tilde{W}_a + C_J\end{aligned} \tag{53}$$

where $\breve{B} = \nabla_2\varphi_c\left(\bar{z}\right)\mathrm{B}\nabla_2\varphi_c(\bar{z})^T$, which is a bounded semi-definite symmetric matrix, $C_J = \bar{\lambda}_{\breve{B}}w_{c0}^2 + \left(2\bar{\lambda}_{\mathrm{B}} + \frac{1}{2}\right)\xi_{d2c0}^2 + \frac{1}{2}w_{c0}^2\varphi_{d2c0}^2 + \hat{w}_{g\max}^2\varphi_{g0}^2 o_{10}^2$. Taylor series for nonlinear term of (52) is utilized.

The third term of (48) is given by

$$\begin{aligned}\dot{\bar{V}}_3 =& -\frac{1}{\alpha_c}\tilde{W}_c^T\dot{\hat{W}}_c \\ =& \tilde{W}_c^T\frac{w}{\left(1+w^Tw\right)^2}\left(-\tilde{W}_c^Tw + \varepsilon_o - \varepsilon_{HJB}\right) \\ &+ 2\tilde{W}_c^T k_s\hat{W}_c + \tilde{W}_c^T k_p\hat{W}_a\end{aligned} \tag{54}$$

in which

$$\begin{aligned}&\tilde{W}_c^T\frac{w}{\left(1+w^Tw\right)^2}\left(-\tilde{W}_c^Tw + \varepsilon_o\right) \\ \le& -\underline{\lambda}_{\mathrm{M}_w}\left\|\tilde{W}_c\right\|^2 + \frac{\varepsilon_o}{\left(1+w^Tw\right)^2}\left(\frac{1}{2}\tilde{W}_c^T\tilde{W}_c + \frac{1}{2}w^Tw\right) \\ \le& -\left(\underline{\lambda}_{\mathrm{M}_w} - \frac{\varepsilon_{o0}}{2}\right)\left\|\tilde{W}_c\right\|^2 + \frac{\varepsilon_{o0}w_{\max}^2}{2}\end{aligned} \tag{55}$$

where $\mathrm{M}_w = \frac{ww^T}{\left(1+w^Tw\right)^2}$ is a bounded semi-definite symmetric matrix since Lemma 9 holds.

$$\begin{aligned}&-\varepsilon_{HJB}\frac{\tilde{W}_c^Tw}{\left(1+w^Tw\right)^2} \\ \le& \frac{1}{\left(1+w^Tw\right)^2}\left(\frac{1}{2}\varepsilon_{HJB}^2 + \frac{1}{2}\left(\tilde{W}_c^Tw\right)^2\right) \\ \le& \varepsilon_{\mathbb{S}} + 2\xi_{d1c0}^2\|K_1\|^2\|\bar{z}_1\|^2 + 2\xi_{d1c0}^2\|\bar{z}_2\|^2 + \frac{1}{2}w_{\max}^2\left\|\tilde{W}_c\right\|^2\end{aligned} \tag{56}$$

where $\varepsilon_{\mathbb{S}} = 2\xi_{d2c0}^2\left\|\bar{f}\left(\bar{z}^+\right) + L\Delta\bar{z}_2^k\right\|^2 + 2\xi_{d1c0}^2L^2\left\|\Delta\bar{z}_1^k\right\|^2$. According to Theorem 1, the signal $\varepsilon_{\mathbb{S}}$ is bounded, i.e. there exists $\varepsilon_{\mathbb{S}0} > 0$, such that $\|\varepsilon_{\mathbb{S}}\| \leqslant \varepsilon_{\mathbb{S}0}$.

$$\begin{aligned}&2\tilde{W}_c^T k_s\hat{W}_c + \tilde{W}_c^T k_p\hat{W}_a \\ \le& \left(\frac{k_p}{2} + k_s\right)W^{*T}W^* - \left(k_s - \frac{k_p}{2}\right)\tilde{W}_c^T\tilde{W}_c - \tilde{W}_c^T k_p\tilde{W}_a\end{aligned} \tag{57}$$

Substituting (55)–(57) into (54), we have

$$\begin{aligned}\dot{\bar{V}}_3 \le& -\kappa_{W_c}\left\|\tilde{W}_c\right\|^2 + C_{W_c} + 2\xi_{d1c0}^2\|K_1\|^2\|\bar{z}_1\|^2 \\ &+ 2\xi_{d1c0}^2\|\bar{z}_2\|^2 - \tilde{W}_c^T k_p\tilde{W}_a\end{aligned} \tag{58}$$

where $\kappa_{W_c} = \underline{\lambda}_{\mathrm{M}_w} + k_s - \frac{\varepsilon_{o0}}{2} - \frac{k_p}{2} - \frac{w_{\max}^2}{2}$, $C_{W_c} = \frac{\varepsilon_{o0}w_{\max}^2}{2} + \varepsilon_{\mathbb{S}0} + \left(k_s + \frac{k_p}{2}\right)w_{c0}^2$.

The fourth term of (48) is given by

$$\begin{aligned}\dot{\bar{V}}_4 =& -\frac{1}{\alpha_a}\tilde{W}_a^T\dot{\hat{W}}_a \\ =& -\tilde{W}_a^T k_p\hat{W}_c + 2k_a\tilde{W}_a^T\hat{W}_a - \tilde{W}_a^T\Upsilon_{W_a}\bar{z}_2 \\ \leqslant& -\bar{\kappa}_{W_a}\left\|\tilde{W}_a\right\|^2 + C_{W_a} + \tilde{W}_c^T k_p\tilde{W}_a - \tilde{W}_a^T\Upsilon_{W_a}\bar{z}_2\end{aligned} \tag{59}$$

where $\bar{\kappa}_{W_a} = k_a - \frac{k_p}{2}$, $C_{W_a} = \left(\frac{k_p}{2} + k_a\right)w_{c0}^2$.

Substituting (51) (53) (58) (59) into (48), we have

$$\begin{aligned}\dot{\bar{V}} \le& -\kappa_{\bar{z}_1}\|\bar{z}_1\|^2 - \kappa_{\bar{z}_2}\|\bar{z}_2\|^2 - \delta e^{-\delta(t-t_k)}\bar{J}_t^* - \kappa_{W_c}\left\|\tilde{W}_c\right\|^2 \\ &- \kappa_{W_a}\left\|\tilde{W}_a\right\|^2 + C_{\bar{z}} + C_J + C_{W_a} + C_{W_c} \\ \le& -\kappa_{\bar{V}}\bar{V} + C_{\bar{V}}\end{aligned} \tag{60}$$

where $\kappa_{\bar{z}_1} = e^{-\delta\Delta t}\underline{\lambda}_{Q_1} + \Pi_1 - 2\xi_{d1c0}^2\|K_1\|^2$, $\kappa_{\bar{z}_2} = e^{-\delta\Delta t}\underline{\lambda}_{Q_2} + \Pi_2 - \frac{1}{2}\bar{\lambda}_{\mathrm{B}} - 2\xi_{d1c0}^2 - \frac{1}{2}$, $\kappa_{W_c} = \underline{\lambda}_{\mathrm{M}_w} + k_s - \frac{\varepsilon_{o0}}{2} - \frac{k_p}{2} - \frac{w_{\max}^2}{2}$, $\kappa_{W_a} = \bar{\kappa}_{W_a} - \bar{\lambda}_{\breve{B}}$, $\kappa_{\bar{V}} = \min\left\{2\kappa_{\bar{z}_1}, 2\kappa_{\bar{z}2}, \delta, 2\alpha_c\kappa_{W_c}, 2\alpha_c\kappa_{W_a}\right\}$, $C_{\bar{V}} = C_{\bar{z}} + C_J + C_{W_c} + C_{W_a}$.

The parameters fulfill that $\kappa_{\bar{z}_1} > 0$, $\kappa_{\bar{z}_2} > 0$, $\delta > 0$, $\kappa_{W_c} > 0$ and $\kappa_{W_a} > 0$. Note that although Π_1 and Π_2 in $\kappa_{\bar{z}_1}$ and $\kappa_{\bar{z}_2}$ are positive definite (from Assumption 5), they cannot be selected directly. So the optimal function of MPC $\bar{J}_t^*$ is added into (47) for introducing $e^{-\delta\Delta t}\underline{\lambda}_{Q_1}$ and $e^{-\delta\Delta t}\underline{\lambda}_{Q_2}$ into $\kappa_{\bar{z}_1}$ and $\kappa_{\bar{z}_2}$, respectively. Then conditions $\kappa_{\bar{z}_1} > 0$ and $\kappa_{\bar{z}_2} > 0$ can be satisfied by choosing appropriate Q_1 and Q_2.

According to Lemma 4, the predictive tracking error $\bar{z}$ and NN's weights errors $\tilde{W}_c$, $\tilde{W}_a$ remain UUB for $t \in [t_k, t_k + T)$, $\forall k \in \mathbb{N}$.

Next the convergence of $\hat{W}_c$ and $\hat{W}_a$ is explained. It can be seen that $\hat{W}_{gk}$ is constant over $t \in [t_k, t_k + T)$, $\forall k \in \mathbb{N}$. if $\hat{W}_{gk} \neq 0$, the signal Υ_{W_a} is persistently existed over $t \in [t_k, t_k + T)$. The last term of (45) ensures sufficient signals of the predictive tracking error space to keep $\hat{W}_a$ from converging to zero, and the last term of (44), which connects the critic network and the action network, keeps $\hat{W}_c$ from converging to zero.

Furthermore, Algorithm 1 solves the optimization problem of MPC iteratively based on the predictive model (10) for $t \in [t_k, t_k + T)$. During each iteration, the initial value of the predictive tracking error, which is gotten from the real system, the updating laws of $\hat{W}_c$, $\hat{W}_a$ and the calculation of $\hat{\bar{\tau}}$ remain the same, except for the initial values of $\hat{W}_c$, $\hat{W}_a$, which are gotten from the previous iteration. Then (60) holds for each iteration. So for each iteration, the closed-loop predictive system is stable with better performance comparing with the previous iteration. Therefore, the following remark can be gotten.

Remark 1. Let Assumptions 3–5, Lemma 8 hold. Under the MPC solving algorithm, the predictive tracking error $\bar{z}$ and NN's weights errors $\tilde{W}_c$, $\tilde{W}_a$ remain UUB. The control strategy $\hat{\bar{\tau}}$ and NN's weights $\hat{W}_c$, $\hat{W}_a$ converge stably to the suboptimal value.

3.3. NN-based MPC for robotic manipulators

Based on the predictive model (10) and MPC solving algorithm, the NN-based MPC strategy is proposed in this section. The suboptimal solution $\hat{\bar{\tau}}$ over $t \in [t_k, t_k + T)$ is solved by Algorithm 1, then it is applied to the robotic manipulator for $t \in [t_k, t_{k+1})$. At time instant t_{k+1}, the parameters $\hat{W}_{f(k+1)}$ and $\hat{W}_{g(k+1)}$ of the predictive model are updated through (11) and (12), the predictive tracking error $\bar{z}\left(t_{k+1}^+\right)$ is updated by real value $z(t_{k+1})$. After that the optimization problem is revisited at time instant t_{k+1} under new initial stations. The architecture of the NN-based MPC strategy is shown in Fig. 1, and the algorithm is summarized in Algorithm 2.

Remark 2. Considering the structure of the control input $\hat{\bar{\tau}}$ in (39), and property of negative definite symmetry for $g(z_1, q_d)$ defined in (6), the initial value of $\hat{W}_{gk}$ should satisfy $\bar{\lambda}\left(\hat{W}_{g0}\varphi_g(z_1(0), q_d(0))\right) \le -\bar{\lambda}\left(M^{-1}(q(0))\right)$. In this initial condition, $\hat{W}_{gk}$ will converge to the neighborhood of W_g^* from the non-zero side, the PE condition of $\hat{\bar{\tau}}$ and the condition $\hat{W}_{gk} \neq 0, \forall k \in \mathbb{N}$ can be guaranteed.

The stability of the closed-loop system is discussed now. An augmented state, which is defined as $\psi \equiv \left[z^T, \bar{z}^T, \tilde{W}_c^T, \tilde{W}_a^T, \Delta\bar{z}^{kT}, \sum_{i=1}^n \tilde{W}_{fk,i}^T, \sum_{i=1}^n \tilde{W}_{gk,i}^T\right]^T$, is adopted to combine all variables. In the following theorem, the NN-based MPC strategy given in Algorithm 2 is shown to guarantee the boundedness of the augmented state ψ.

Theorem 3. *For the robotic manipulator (3) with input constraints (4), let Assumptions 1–5, Lemmas 1–8 and Remark 2 hold. Under the NN-based MPC strategy given in Algorithm 2, the augmented state ψ remains UUB, namely the robot tracking error z, the predictive tracking error $\bar{z}$, the estimation error $\Delta\bar{z}^k$ and NN's weights errors $\tilde{W}_{fk}$, $\tilde{W}_{gk}$, $\tilde{W}_c$, $\tilde{W}_a$ remain UUB, if all conditions in Theorems 1 and 2 hold. z, $\tilde{W}_{fk}$, $\tilde{W}_{gk}$, $\tilde{W}_c$ and $\tilde{W}_a$ will converge to the compact sets Ω_z, Ω_{W_f}, Ω_{W_g}, Ω_{W_c}, Ω_{W_a}:*

$$\Omega_z := \left\{z \in \Re^{2n} \middle| \|z\| \le \sqrt{M}\right\}$$

$$\Omega_{W_f} := \left\{\tilde{W}_{fk} \in \Re^{l_f \times n} \middle| \left\|\tilde{W}_{fk}\right\| \le \sqrt{2\alpha_f M}\right\}$$

$$\Omega_{W_g} := \left\{\tilde{W}_{gk} \in \Re^{l_g \times n} \middle| \left\|\tilde{W}_{gk}\right\| \le \sqrt{2\alpha_g M}\right\}$$

$$\Omega_{W_c} := \left\{\tilde{W}_c \in \Re^{l_c \times 1} \middle| \left\|\tilde{W}_c\right\| \le \sqrt{2\alpha_c M}\right\}$$

$$\Omega_{W_a} := \left\{\tilde{W}_a \in \Re^{l_c \times 1} \middle| \left\|\tilde{W}_a\right\| \le \sqrt{2\alpha_a M}\right\}$$

where $M = 3\mathbb{Z}_{\bar{V}} + \frac{3}{2}\mathbb{Z}$, $\mathbb{Z}_{\bar{V}}$ will be defined later.

Algorithm 2 NN-based MPC strategy for robot tracking control

Input:
- $\hat{W}_{f0}$, $\hat{W}_{g0}$: Initial NN's weights of the predictive model;
- T: The prediction horizon;
- Δt: The solving interval of the optimization problem for the NN-based MPC;
- α_f, α_g, k_f, k_g: Learning rates/parameters of NN for the predictive model;

Initialize $q_1(0)$, $q_2(0)$, $k = 0$, $t_0 = 0$;
Compute $z(0)$ according to (5);
repeat
  if $t = t_k$ **then**
    Solve the MPC problem via Algorithm 1 for suboptimal control strategy $\hat{\bar{\tau}}$;
  end if
  Compute $t_{k+1} = t_k + \Delta_t$;
  for $t \in [t_k, t_{k+1})$ **do**
    Apply $\hat{\bar{\tau}}$ to the robotic manipulator and observe q, $\dot{q}$;
    Compute $z(t)$ according to (5);
  end for
  if $t = t_{k+1}$ **then**
    Compute the estimation error of tracking error via $\Delta\bar{z}_j^{k+1} = z_j(t_{k+1}) - \bar{z}_j\left(t_{k+1}^-\right) \quad j = 1, 2$;
    Update NN's weights $\hat{W}_{f(k+1)}$, $\hat{W}_{g(k+1)}$ according to (11) and (12);
    Update the predictive tracking error via $\bar{z}\left(t_{k+1}^+\right) = z(t_{k+1})$;
  end if
  $k = k + 1$;
until the end of control period

Proof. Construct a Lyapunov function candidate as

$$V = V_{mk} + \bar{V} + V_z \tag{61}$$

where $V_z = \frac{1}{2}z^T z$.

From Theorem 1, it can be gotten that the estimation error $\Delta\bar{z}^k$ and NN's weights errors $\tilde{W}_{fk}$, $\tilde{W}_{gk}$ remain UUB. V_{mk} satisfies $V_{mk}(\infty) \le \frac{1}{2}\mathbb{Z}$.

Then for the predictive tracking error $\bar{z}$ and NN's weights errors $\tilde{W}_c$, $\tilde{W}_a$, mathematical induction (MI) is utilized to explain that they are UUB for the whole control period.

Firstly, for $k = 0$, $t \in (t_0, t_1)$, Multiplying (60) by $e^{\kappa_{\bar{V}} t}$ and integrating the inequality, we have

$$\bar{V}(t) \le e^{-\kappa_{\bar{V}} t}\bar{V}(t_0) + \frac{C_{\bar{V}}}{\kappa_{\bar{V}}} - \frac{C_{\bar{V}}}{\kappa_{\bar{V}}}e^{-\kappa_{\bar{V}} t} \tag{62}$$

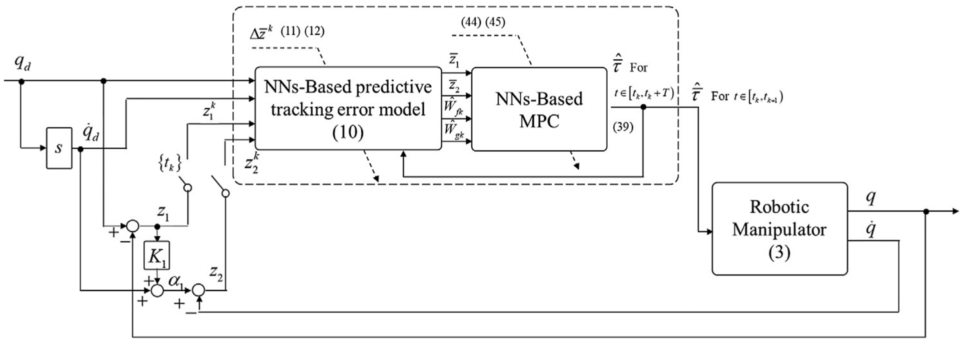


Fig. 1. Architecture of the NN-based model predictive tracking control.

At time instant t_1, the predictive tracking error $\bar z$ is revised by real value through $\bar z\left(t_1^+\right)=z\left(t_1\right)$. For solving $\bar V\left(t_1^+\right)$, we have the following assumption holds:

Assumption 6. The function $\bar J_t^*\left(\bar z\right)\ :\ \Re^n\rightarrow\Re^1$ is a Lipschitz continuity mapping, which means

$$\bar J_t^*\left(\bar z\left(t_1^+\right)\right)-\bar J_t^*\left(\bar z\left(t_1^-\right)\right)\leq L_{\bar J}\left\|\Delta\bar z^1\right\| \tag{63}$$

Considering $\frac{1}{2}\bar z^T\left(t_1^+\right)\bar z\left(t_1^+\right)=\frac{1}{2}\bar z^T\left(t_1^-\right)\bar z\left(t_1^-\right)+\frac{1}{2}\Delta\bar z^{1T}\Delta\bar z^1+\bar z\left(t_1^-\right)^T\Delta\bar z^1$, we have

$$\begin{aligned}\bar V\left(t_1^+\right)=&\bar V\left(t_1^-\right)+\frac{1}{2}\Delta\bar z^{1T}\Delta\bar z^1+\bar z\left(t_1^-\right)^T\Delta\bar z^1\\&+e^{-\delta\Delta t}L_{\bar J}\left\|\Delta\bar z^1\right\|\end{aligned} \tag{64}$$

According to Theorems 1 and 2, we can get that $\Delta\bar z^k$ and $\bar z\left(t_k^-\right)$ are bounded, i.e. there exist $\sigma_\Delta>0$, $\sigma_{\bar z}>0$ such that $\left\|\Delta\bar z^k\right\|\leq\sigma_\Delta$, $\left\|\bar z\left(t_k^-\right)\right\|\leq\sigma_{\bar z}$. Introducing (62), (64) becomes

$$\bar V\left(t_1^+\right)\leq e^{-\kappa_{\bar V}t_1}\bar V\left(t_0\right)+\frac{C_{\bar V}}{\kappa_{\bar V}}-\frac{C_{\bar V}}{\kappa_{\bar V}}e^{-\kappa_{\bar V}t_1}+\sigma \tag{65}$$

where $\sigma=\frac{1}{2}\sigma_\Delta^2+\left(e^{-\delta\Delta t}L_{\bar J}+\sigma_{\bar z}\right)\sigma_\Delta$.

Secondly, for $k=1$, $t\in\left(t_1,t_2\right)$, similar to the first step, we have

$$\begin{aligned}\bar V\left(t\right)&\leq e^{-\kappa_{\bar V}t}e^{\kappa_{\bar V}t_1}\bar V\left(t_1^+\right)+\frac{C_{\bar V}}{\kappa_{\bar V}}-\frac{C_{\bar V}}{\kappa_{\bar V}}e^{\kappa_{\bar V}t_1}e^{-\kappa_{\bar V}t}\\&\leq e^{-\kappa_{\bar V}t}\bar V\left(t_0\right)+\frac{C_{\bar V}}{\kappa_{\bar V}}-\frac{C_{\bar V}}{\kappa_{\bar V}}e^{-\kappa_{\bar V}t}+e^{-\kappa_{\bar V}\left(t-t_1\right)}\sigma\end{aligned} \tag{66}$$

$$\bar V\left(t_2^+\right)\leq e^{-\kappa_{\bar V}t}\bar V\left(t_0\right)+\frac{C_{\bar V}}{\kappa_{\bar V}}-\frac{C_{\bar V}}{\kappa_{\bar V}}e^{-\kappa_{\bar V}t_2}+e^{-\kappa_{\bar V}\Delta t}\sigma+\sigma \tag{67}$$

Then, suppose for $t\in\left(t_k,t_{k+1}\right),k=1,2,3\ldots$, following conditions hold:

$$\begin{aligned}\bar V\left(t\right)\leq&e^{-\kappa_{\bar V}t}\bar V\left(t_0\right)+\frac{C_{\bar V}}{\kappa_{\bar V}}-\frac{C_{\bar V}}{\kappa_{\bar V}}e^{-\kappa_{\bar V}t}\\&+e^{-\kappa_{\bar V}\left(t-t_k\right)}\sigma\sum_{i=0}^{k-1}e^{-i\kappa_{\bar V}\Delta t}\end{aligned} \tag{68}$$

$$\begin{aligned}\bar V\left(t_{k+1}^+\right)\leq&e^{-\kappa_{\bar V}t_{k+1}}\bar V\left(t_0\right)+\frac{C_{\bar V}}{\kappa_{\bar V}}-\frac{C_{\bar V}}{\kappa_{\bar V}}e^{-\kappa_{\bar V}t_{k+1}}\\&+\sigma\sum_{i=0}^{k}e^{-i\kappa_{\bar V}\Delta t}\end{aligned} \tag{69}$$

For $t\in\left(t_{k+1},t_{k+2}\right)$, it can be gotten

$$\begin{aligned}\bar V\left(t\right)\leq&e^{-\kappa_{\bar V}t}e^{\kappa_{\bar V}t_{k+1}}\bar V\left(t_{k+1}^+\right)+\frac{C_{\bar V}}{\kappa_{\bar V}}-\frac{C_{\bar V}}{\kappa_{\bar V}}e^{\kappa_{\bar V}t_{k+1}}e^{-\kappa_{\bar V}t}\\\leq&e^{-\kappa_{\bar V}t}\bar V\left(t_0\right)+\frac{C_{\bar V}}{\kappa_{\bar V}}-\frac{C_{\bar V}}{\kappa_{\bar V}}e^{-\kappa_{\bar V}t}\\&+e^{-\kappa_{\bar V}\left(t-t_{k+1}\right)}\sigma\sum_{i=0}^{k}e^{-i\kappa_{\bar V}\Delta t}\end{aligned} \tag{70}$$

$$\begin{aligned}\bar V\left(t_{k+2}^+\right)\leq&\bar V\left(t_{k+2}^-\right)+\sigma\\\leq&e^{-\kappa_{\bar V}t_{k+2}}\bar V\left(t_0\right)+\frac{C_{\bar V}}{\kappa_{\bar V}}-\frac{C_{\bar V}}{\kappa_{\bar V}}e^{-\kappa_{\bar V}t_{k+2}}+\sigma\sum_{i=0}^{k+1}e^{-i\kappa_{\bar V}\Delta t}\end{aligned} \tag{71}$$

It is obvious that $\forall t>0$, the following condition holds:

$$\bar V\left(t\right)\leq e^{-\kappa_{\bar V}t}\bar V\left(t_0\right)+\frac{C_{\bar V}}{\kappa_{\bar V}}-\frac{C_{\bar V}}{\kappa_{\bar V}}e^{-\kappa_{\bar V}t}+\frac{\sigma}{1-e^{-\kappa_{\bar V}\Delta t}} \tag{72}$$

as $\lim_{k\to\infty}\sum_{i=0}^{k}e^{-i\kappa_{\bar V}\Delta t}=\frac{1}{1-e^{-\kappa_{\bar V}\Delta t}}$. Then it can be gotten that $\bar V\left(\infty\right)\leq\mathbb{Z}_{\bar V}$, where $\mathbb{Z}_{\bar V}=\frac{C_{\bar V}}{\kappa_{\bar V}}+\frac{\sigma^*}{1-e^{-\kappa_{\bar V}\Delta t}}$.

About the robot tracking error z, for $t\in\left[t_k,t_{k+1}\right),\forall k\in\mathbb{N}$, we have

$$V_z=\frac{1}{2}\left(\bar z+\Delta\bar z\right)^T\left(\bar z+\Delta\bar z\right)\leq\bar z^T\bar z+\Delta\bar z^T\Delta\bar z \tag{73}$$

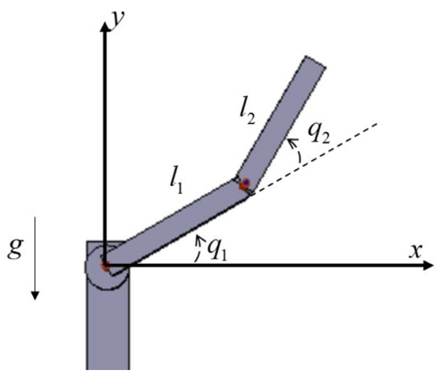
where $\Delta\bar z\left(t\right)=z\left(t\right)-\bar z\left(t\right)$.

It is worth noting that $\Delta\bar z$ has the same convergence with $\Delta\bar z^k$, which means $\Delta\bar z^T\left(\infty\right)\Delta\bar z\left(\infty\right)\leq\mathbb{Z}$. Then it can be gotten that $V_z\left(\infty\right)\leqslant2\mathbb{Z}_{\bar V}+\mathbb{Z}$.

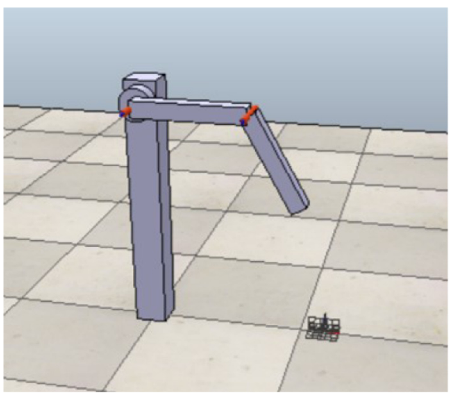
From above analysis, we can conclude that $V\left(\infty\right)\leqslant3\mathbb{Z}_{\bar V}+\frac{3}{2}\mathbb{Z}$. Then Theorem 3 is proved, the augmented state ψ remains UUB under the proposed NN-based MPC strategy.

4. Simulation studies

To demonstrate the effectiveness of the proposed NN-based MPC strategy, co-simulation based on CoppeliaSim (V-REP) and Matlab for a 2-DOF robotic manipulator is given in this section. The robotic system is defined by (3) with parameters given in Appendix. The model of the robotic manipulator established in CoppeliaSim is shown in Fig. 2. The input constraints are described as $\left|\tau_i\left(t\right)\right|\leqslant12,i=1,2$. The desired trajectory is defined by $q_d\left(t\right)=\left[q_{1d},q_{2d}\right]^T=\left[\frac{\pi}{4}\cos\left(\frac{\pi}{4}t\right),\frac{\pi}{2}\sin\left(\frac{\pi}{4}t\right)\right]^T$. The initial condition of the robotic manipulator is chosen as $q\left(0\right)=\left[0,-\frac{\pi}{3}\right]^T$, $\dot q\left(0\right)=\left[0,0\right]^T$. The whole simulation time is chosen as $T_s=20$ s.



(a) Coordinate of the robotic manipulator.



(b) Model of the robotic manipulator in CoppeliaSim.

Fig. 2. Model of the robotic manipulator.

Comparative discussions are accomplished with the Lyapunov function method (LFM) [2,41], the constrained MPC (CMPC) strategy [8] and the strategy combined MPC and integral sliding mode controller (MPC-ISM) [21]. The algorithms run at a laptop (Intel(R) Core(TM) i5-8265U @1.60 GHz). The simulation environments are chosen as CoppeliaSim Edu V4.0.0 rev4 and MATLAB2019b. Simulation details and results are shown as follows.

4.1. Simulation description

(a) The Lyapunov function method

The Lyapunov function method is taken into account firstly. Referring to [2,41], the controller can be designed with a function $S(\tau_L)$ as

$$S(\tau_{Li}) = \begin{cases} \lambda sign(\tau_{Li}) & |\tau_{Li}| > \lambda \\ \tau_{Li} & |\tau_{Li}| \le \lambda \end{cases}$$

where $\tau_L = -z_1 + K_p(z_2 + \xi) + \hat{W}^T\varphi(z_L)$. The parameter of the auxiliary variable in (5) is defined as $K_1 = 2$. The control gain matrix is chosen as $K_p = \mathrm{diag}(5, 0.5)$. The auxiliary variable ξ is used for reducing the input constraints effects. The updating law and corresponding parameters of ξ can refer to [41]. $\hat{W}^T\varphi(z_L)$ is a neural network which is used for estimating $M(q)\dot{\alpha}_1 + C(q, \dot{q})\alpha_1 + G(q)$. The activation function $\varphi(z_L)$ is chosen as Gaussian with input signal $z_L = \left[q^T, \dot{q}^T, \alpha_1^T, \dot{\alpha}_1^T\right]$. The adaptive law of NN is designed as $\dot{\hat{W}}_k = \Gamma\left(\varphi(z_L)z_{2k} - \sigma\hat{W}_k\right)$ with parameters $\Gamma = 0.06$ and $\sigma = 0.1$. The joints angle tracking and control torques are shown in Figs. 3 and 5(a). Comparative results are discussed later.

(b) The NN-Based model predictive control

In this part, we discuss the simulation under the proposed NN-Based MPC strategy. Two groups of NNs have been developed under the proposed MPC structure. In the part of the predictive model, NNs with 64 and 36 hidden-layer-nodes are used for $\hat{W}_{fk}\varphi_f\left(\bar{z}^+\right)$ and $\hat{W}_{gk}\varphi_g(\bar{z}_1, q_d)$, respectively. The centers for $\varphi_f\left(\bar{z}^+\right)$ and $\varphi_g(\bar{z}_1, q_d)$ are chosen in the area of $[-1, 1]\times[-1, 1]\times[-1, 1]\times[-1, 1]\times[-1.6, 1.6]\times[-1.6, 1.6]\times[0]\times[0]\times[0]\times[0]$ and $[-2, 2]\times[-1, 1]\times[-1.6, 1.6]\times[-1.6, 1.6]$. The variance is set to be 25. Considering that $g(z_1, q_d)$ is a negative symmetric definite matrix, its estimation is expressed as $\hat{W}_{gk}\varphi_g(\bar{z}_1, q_d) = \begin{bmatrix}\hat{W}_{gk,1} & \hat{W}_{gk,2} \\ \hat{W}_{gk,2} & \hat{W}_{gk,3}\end{bmatrix}\varphi_g(\bar{z}_1, q_d)$. The parameter L in (10) is chosen as $L = 0.35$. The parameter of the auxiliary variable in (5) is defined as $K_1 = 5$.

In the part of solving the optimization problem of MPC, parameters of the cost function are chosen as $Q_1 = \mathrm{diag}(200, 200)$, $Q_2 = \mathrm{diag}(5, 5)$, $R = \mathrm{diag}\left(\frac{1}{5}, \frac{1}{2}\right)$. For ensuring real-time performance and control accuracy, the solving interval of the optimization problem and the prediction horizon are chosen as $\Delta t = 0.05$ s and $T = 0.07$ s, respectively. NNs with 81 hidden-layer-nodes are used for $\hat{W}_c\varphi_c(\bar{z})$ and $\hat{W}_a\varphi_c(\bar{z})$. The centers for $\varphi_c(\bar{z})$ are chosen in the area of $[-2, 0, 2]\times[-1, 0, 1]\times[-1, 0, 1]\times[-1, 0, 1]$. The variance is set to be 25, too. The joints angle tracking and control torques are shown in Figs. 3 and 5(b).

For further illustrating the effectiveness of the proposed control strategy, we consider the comparison with the constrained MPC strategy and the strategy combined MPC and integral sliding mode controller in next parts.

(c) The CMPC strategy

For the CMPC strategy, parameters of the cost function are chosen as same as the NN-based MPC. Quadratic form $\bar{z}^TQ\bar{z}$ is used for the terminal penalty $\Psi(\bar{z}(t_k + T))$, in which $Q = \mathrm{diag}(Q_1, Q_2)$. Parameters of the robotic manipulator are assumed known imprecisely, for example, there is an error of 0.005 kg in the masses of link 1 and link 2, and an error of 0.0005 m in the lengths. The optimization problem is solved by the Gurobi solver. The joints angle tracking and control torques are shown in Figs. 3 and 5(c).

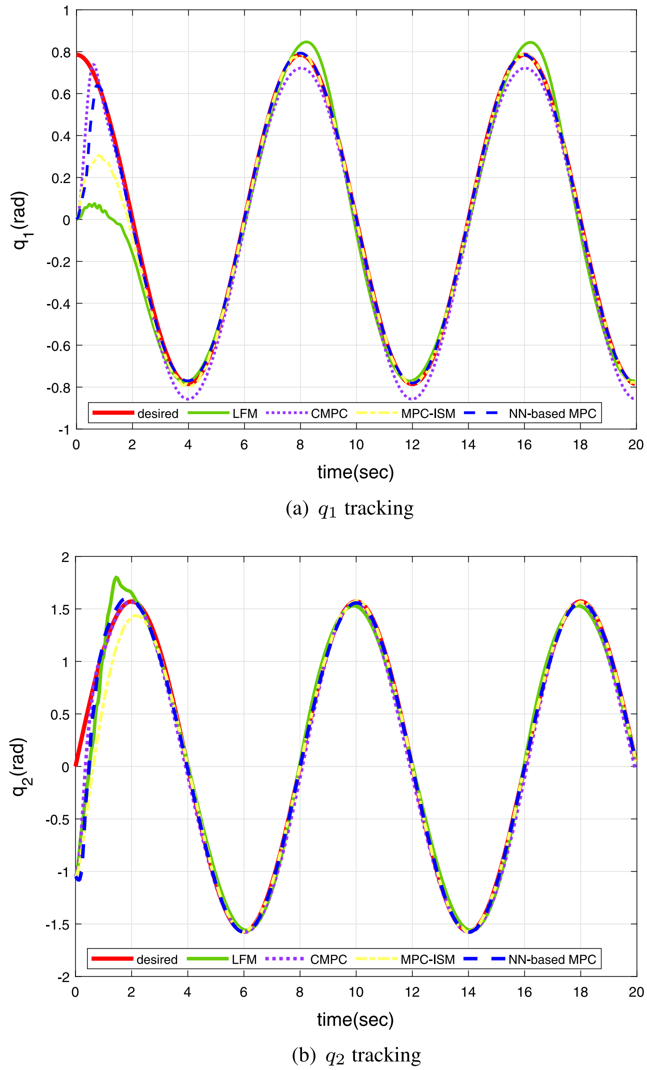
(d) The MPC-ISM strategy

For the MPC-ISM strategy, the control parameters are the same as those in [21]. Assumptions about the parameters of the robotic manipulator are the same as the CMPC strategy. The optimization problem is solved by the Gurobi solver, too. The joints angle tracking and control torques are shown in Figs. 3 and 5(d).

4.2. Results and discussion

Comparisons of joints tracking errors with four control strategies are shown in Fig. 4. It can be seen that results with all four kinds of control schemes are convergent, but the tracking errors with the Lyapunov function method and CMPC strategy are much larger than those with the NN-based MPC strategy. Specifically, (1) for the Lyapunov function method under selected control parameters, q_1 converges slower than the other methods, while q_2 has obvious overshoots. (2) For the CMPC method, there exist steady-state errors under the influence of model uncertainties. (3) For the MPC-ISM scheme, both q_1 and q_2 have a slower initial response. There also exist chattering phenomenons of q_1 and q_2 because of the introduction of the ISM controller. (4) For the NN-based MPC strategy, as adaptive NN is adopted to compensate for model uncertainties, and predictive control strategy is used to calculate the optimal control law, the good tracking performance can be guaranteed.

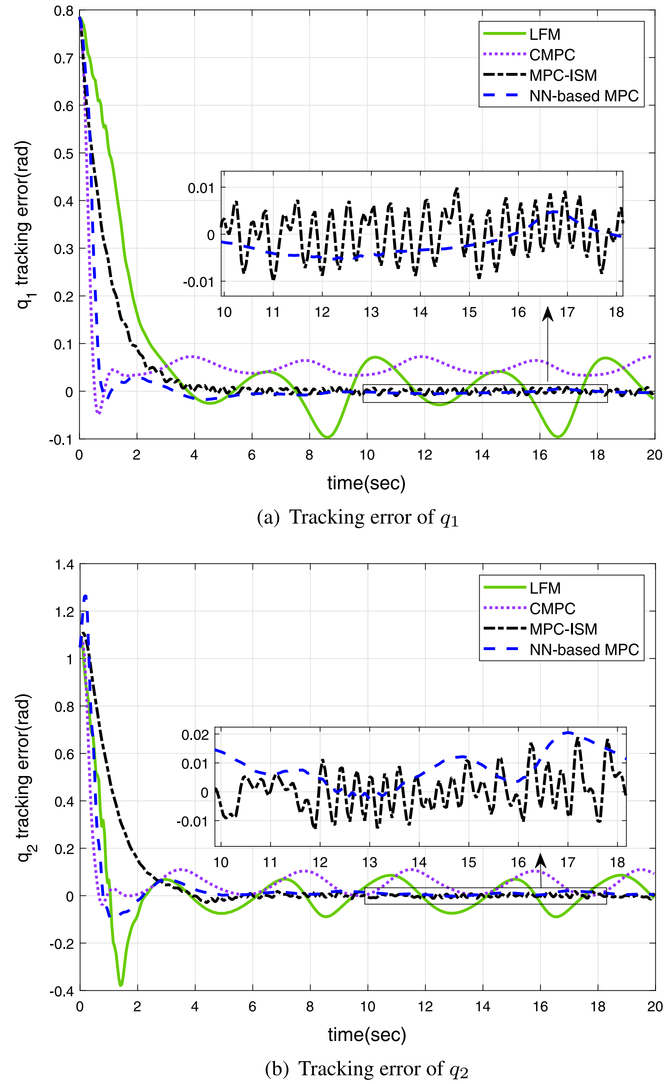
Furthermore, for CMPC and MPC-ISM schemes, the parameters $M(q)$, $C(q, \dot{q})$ and $G(q)$ in (3) are used directly for control design.



(a) q_1 tracking

(b) q_2 tracking

Fig. 3. Joints tracking with different control strategies.



(a) Tracking error of q_1

(b) Tracking error of q_2

Fig. 4. Joints tracking errors with different control strategies.

So the expandability of these methods is worse than the other two methods.

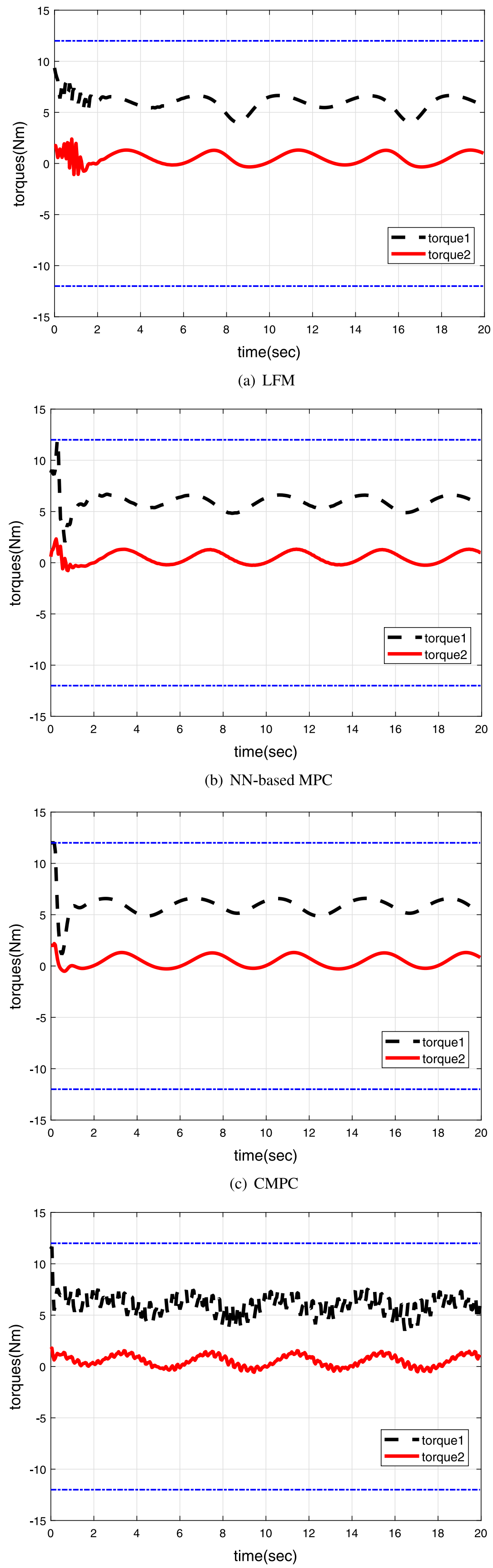
In addition, Fig. 5 shows that control torques with all four kinds of control schemes satisfy constraints. But control torques with the Lyapunov function method are volatile at the beginning of the control period. For the MPC-ISM strategy, there exist chattering phenomenons of torques in the whole control period. It can be gotten from Fig. 5 that control torques with the other two methods are more appropriate to the real system.

We have verified the favorable tracking capability and feasible control torques of the proposed NN-based MPC strategy. Next, its computation burden will be discussed and analyzed. Firstly, in this paper, the discrete updating mode of NNs in the predictive model can reduce the computation burden of NN updating, comparing with the continuous updating mode of NN in the Lyapunov function method. Then the prediction horizon is chosen as $T = 1.4\Delta t$ to reduce the solving time of optimization problem and keep control performance at the same time. Fig. 6 shows the calculation time of solving the optimization problem at time instant t_k, which is much smaller than the solving interval Δt. Physical simulation results with CoppeliaSim also show that the whole simulation calculation time for the NN-based MPC strategy is not greater than the actual time. These results indicate that the calculation burden is acceptable for the real-time implementation.

In conclusion, the NN-based MPC strategy for robotic manipulators proposed in this paper can achieve competitive performance in handling the unknown dynamics with input constraints.

5. Conclusions

In this paper, an NN-based MPC strategy was developed for robotic manipulators with unknown dynamics and input constraints. The proposed structure contained two groups of NNs. The first group of NNs was adopted as a predictive model of MPC for the robotic system. Online learning strategies, which were based on errors between predictive tracking error and the actual one, were established to handle the robotic unknown dynamics. Based on the predictive model, the second group of NNs was applied to solve the optimization problem of MPC. An actor–critic scheme with different weights and the same activation function was adopted, and adaptive learning strategies were established for balancing between optimal tracking performance and predictive system stability. A nonquadratic cost function was developed for handling the input constraints. According to the



(a) LFM

(b) NN-based MPC

(c) CMPC

(d) MPC-ISM

Fig. 5. Control torques.

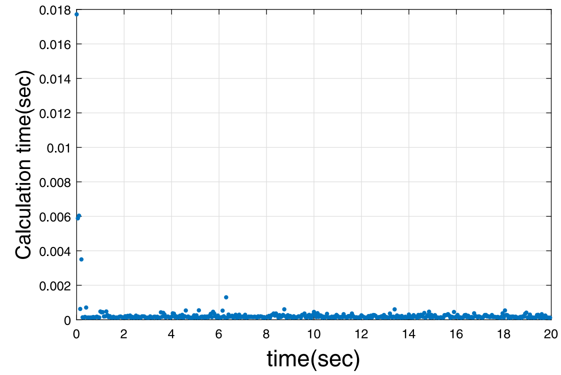


Fig. 6. Calculation time of solving the optimization problem at time instant t_k.

Lyapunov theorem, it was proved that all variables of the closed-loop system were UUB under the desired strategy. Simulation studies were carried out to illustrate the effectiveness of the proposed control strategy, comparing with the Lyapunov function method, the CMPC strategy and the MPC-ISM method.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported in part by the National Key Research and Development Program of China (Grants No. 2017YFB1300200, 2017YFB1300203), the National Natural Science Foundation of China (Grants No. 91648205, 61627808 and 91948303), the Strategic Priority Research Program of Chinese Academy of Science (Grant No. XDB32050100) and the development of science and technology of Guangdong province special fund project, China (Grant No. 2016B090910001).

Appendix

A 2-DOF robotic manipulator defined by (3) is adopted for demonstrating the effectiveness of the proposed method. The inertia matrix $M(q)$, Centripetal and Coriolis force $C(q, \dot{q})$ and gravitational force $G(q)$ are defined as

$$M(q) = \begin{bmatrix} a_1 + a_2 + 2a_3 \cos q_2 & a_2 + a_3 \cos q_2 \\ a_2 + a_3 \cos q_2 & a_2 \end{bmatrix}$$

$$C(q, \dot{q}) = \begin{bmatrix} -a_3 \dot{q}_2 \sin q_2 & -a_3 (\dot{q}_1 + \dot{q}_2) \sin q_2 \\ a_3 \dot{q}_1 \sin q_2 & 0 \end{bmatrix}$$

$$G(q) = \begin{bmatrix} a_4 g \cos q_1 + a_5 g \cos (q_1 + q_2) \\ a_5 g \cos (q_1 + q_2) \end{bmatrix}$$

where $a_1 = m_1 l_{c1}^2 + m_2 l_1^2 + I_1$, $a_2 = m_2 l_{c2}^2 + I_2$, $a_3 = m_2 l_1 l_{c2}$, $a_4 = m1 l_{c2} + m_2 l_1$, $a_5 = m_2 l_{c2}$. l_i and m_i are the length and mass of link i, l_{ci} is the distance from joint $i-1$ to the center of mass of link i, I_i is the moment of inertia of link i about the axis perpendicular to the plane of link and passing through the center of mass of link i.

The value of parameters refer to [41]. They are given as follows: $m_1 = 2.0$ kg, $m_2 = 0.85$ kg, $l_1 = 0.35$ m, $l_2 = 0.31$ m, $l_{c1} = 0.175$ m, $l_{c2} = 0.155$ m, $I_1 = 61.25 \times 10^{-3}$ kg m^2, $I_2 = 20.42 \times 10^{-3}$ kg m^2.

References

[1] Jin L, Li S, Yu J, He J. Robot manipulator control using neural networks: A survey. Neurocomputing 2018;285:23–34. http://dx.doi.org/10.1016/j.neucom.2018.01.002.

[2] He W, Chen Y, Yin Z. Adaptive neural network control of an uncertain robot with full-state constraints. IEEE Trans Cybern 2016;46(3):620–9. http://dx.doi.org/10.1109/tcyb.2015.2411285.

[3] Islam S, Liu X. Robust sliding mode control for robot manipulators. IEEE Trans Ind Electron 2011;58(6):2444–53. http://dx.doi.org/10.1109/tie.2010.2062472.

[4] Rahmani M, Ghanbari A, Ettefagh MM. Hybrid neural network fraction integral terminal sliding mode control of an inchworm robot manipulator. Mech Syst Signal Process 2016;80:117–36. http://dx.doi.org/10.1016/j.ymssp.2016.04.004.

[5] He W, Dong Y. Adaptive fuzzy neural network control for a constrained robot using impedance learning. IEEE Trans Neural Netw Learn Syst 2018;29(4):1174–86. http://dx.doi.org/10.1109/TNNLS.2017.2665581.

[6] Ma C, Qiao H. Distributed asynchronous event-triggered consensus of nonlinear multi-agent systems with disturbances: An extended dissipative approach. Neurocomputing 2017;243:103–14. http://dx.doi.org/10.1016/j.neucom.2017.03.005.

[7] Ma C, Shi P, Zhao X, Zeng Q. Consensus of euler–Lagrange systems networked by sampled-data information with probabilistic time delays. IEEE Trans Cybern 2015;45(6):1126–33. http://dx.doi.org/10.1109/TCYB.2014.2345735.

[8] Mayne DQ, Rawlings J, Rao C, Scokaert P. Constrained model predictive control: Stability and optimality. Automatica 2000;36(6):789–814. http://dx.doi.org/10.1016/s0005-1098(99)00214-9.

[9] Wang T, Gao H, Qiu J. A combined adaptive neural network and nonlinear model predictive control for multirate networked industrial process control. IEEE Trans Neural Netw Learn Syst 2016;27(2):416–25. http://dx.doi.org/10.1109/tnnls.2015.2411671.

[10] Donoso F, Mora A, Cardenas R, Angulo A, Saez D, Rivera M. Finite-set model predictive control strategies for a 3L-NPC inverter operating with fixed switching frequency. IEEE Trans Ind Electron 2018;65(5):3954–65. http://dx.doi.org/10.1109/TIE.2017.2760840.

[11] Halvgaard R, Vandenberghe L, Poulsen K, Niels, Madsen H, Jrgensen B. Economic model predictive control for smart energy systems. IEEE Trans Smart Grid 2016;7(3):1675–82. http://dx.doi.org/10.1109/TSG.2016.2526077.

[12] Justo JJ, Mwasilu F, Kim E-K, Kim J, Choi HH, Jung J-W. Fuzzy model predictive direct torque control of IPMSMs for electric vehicle applications. IEEE/ASME Trans Mechatronics 2017;22(4):1542–53. http://dx.doi.org/10.1109/TMECH.2017.2665670.

[13] Worthmann K, Mehrez MW, Zanon M, Mann GKI, Gosine RG, Diehl M. Model predictive control of nonholonomic mobile robots without stabilizing constraints and costs. IEEE Trans Control Syst Technol 2016;24(4):1394–406. http://dx.doi.org/10.1109/TCST.2015.2488589.

[14] Shen C, Shi Y, Buckham B. Trajectory tracking control of an autonomous underwater vehicle using Lyapunov-based model predictive control. IEEE Trans Ind Electron 2018;65(7):5796–805. http://dx.doi.org/10.1109/TIE.2017.2779442.

[15] Kayacan E, Ramon H, Saeys W. Robust trajectory tracking error-based model predictive control for unmanned ground vehicles. IEEE/ASME Trans Mechatronics 2016;21(2):806–14. http://dx.doi.org/10.1109/TMECH.2015.2492984.

[16] Sun Z, Li D, Xia Y, Liu K. Event-based model predictive tracking control of nonholonomic systems with coupled input constraint and bounded disturbances. IEEE Trans Automat Control 2018;63(2):608–15. http://dx.doi.org/10.1109/TAC.2017.2736518.

[17] Liu C, Gao J, Li H, Xu D. Aperiodic robust model predictive control for constrained continuous-time nonlinear systems: An event-triggered approach. IEEE Trans Cybern 2018;48(5):1397–405. http://dx.doi.org/10.1109/TCYB.2017.2695499.

[18] Farina M, Scattolini R. Tube-based robust sampled-data MPC for linear continuous-time systems. Automatica 2012;48(7):1473—1476. http://dx.doi.org/10.1016/j.automatica.2012.03.026.

[19] Falugi P, Mayne DQ. Getting robustness against unstructured uncertainty: A tube-based MPC approach. IEEE Trans Automat Control 2014;59(5):1290–5. http://dx.doi.org/10.1109/TAC.2013.2287727.

[20] Zhang R, Wu S, Cao Z, Lu J, Gao F. A systematic min-max optimization design of constrained model predictive tracking control for industrial processes against uncertainty. IEEE Trans Control Syst Technol 2018;26(6):2157–64. http://dx.doi.org/10.1109/TCST.2017.2748059.

[21] Incremona GP, Ferrara A, Magni L. MPC for robot manipulators with integral sliding modes generation. IEEE/ASME Trans Mechatronics 2017;22(3):1299–307. http://dx.doi.org/10.1109/TMECH.2017.2674701.

[22] Faulwasser T, Weber T, Zometa P, Findeisen R. Implementation of nonlinear model predictive path-following control for an industrial robot. IEEE Trans Control Syst Technol 2017;25(4):1505–11. http://dx.doi.org/10.1109/TCST.2016.2601624.

[23] Dong L, Yan J, Yuan X, He H, Sun C. Functional nonlinear model predictive control based on adaptive dynamic programming. IEEE Trans Cybern 2019;49(12):4206–18. http://dx.doi.org/10.1109/TCYB.2018.2859801.

[24] Patan K. Neural network-based model predictive control: Fault tolerance and stability. IEEE Trans Control Syst Technol 2015;23(3):1147–55. http://dx.doi.org/10.1109/TCST.2014.2354981.

[25] Patan K. Two stage neural network modelling for robust model predictive control. ISA Trans 2018;72:56–65. http://dx.doi.org/10.1016/j.isatra.2017.10.011.

[26] Pan Y, Wang J. Model predictive control of unknown nonlinear dynamical systems based on recurrent neural networks. IEEE Trans Ind Electron 2012;59(8):3089–101. http://dx.doi.org/10.1109/tie.2011.2169636.

[27] Lian C, Xu X, Chen H, He H. Near-optimal tracking control of mobile robots via receding-horizon dual heuristic programming. IEEE Trans Cybern 2016;46(11):2484–96. http://dx.doi.org/10.1109/TCYB.2015.2478857.

[28] Yan Z, Wang J. Robust model predictive control of nonlinear systems with unmodeled dynamics and bounded uncertainties based on neural networks. IEEE Trans Neural Netw Learn Syst 2014;25(3):457–69. http://dx.doi.org/10.1109/TNNLS.2013.2275948.

[29] Akpan VA, Hassapis GD. Nonlinear model identification and adaptive model predictive control using neural networks. ISA Trans 2011;50(2):177–94. http://dx.doi.org/10.1016/j.isatra.2010.12.007.

[30] Yang W, Zhang W, Xu D, Yan W. Fuzzy model predictive control for 2-DOF robotic arms. Assem Autom 2018;38(5):568–75. http://dx.doi.org/10.1108/AA-11-2017-162.

[31] Carron A, Arcari E, Wermelinger M, Hewing L, Hutter M, Zeilinger MN. Data-driven model predictive control for trajectory tracking with a robotic arm. IEEE Robot Autom Lett 2019;4(4):3758–65. http://dx.doi.org/10.1109/LRA.2019.2929987.

[32] Wu BF, Lin C-H. Adaptive neural predictive control for permanent magnet synchronous motor systems with long delay time. IEEE Access 2019;7:108061–9. http://dx.doi.org/10.1109/ACCESS.2019.2932746.

[33] Chen Y, Li Z, Kong H, Ke F. Model predictive tracking control of nonholonomic mobile robots with coupled input constraints and unknown dynamics. IEEE Trans Ind Inf 2019;15(6):3196–205. http://dx.doi.org/10.1109/TII.2018.2874182.

[34] Farrokhi M, Mazdarani H. Adaptive neuro-predictive position/velocity control of robot manipulators in work space. In: International Conference on Methods and MODELS in Automation and Robotics. IEEE; 2012, p. 349–54. http://dx.doi.org/10.1109/MMAR.2012.6347864.

[35] Hajiloo A, Keshmiri M, Xie WF, Wang TT. Robust on-line model predictive control for a constrained image based visual servoing. IEEE Trans Ind Electron 2016;63(4):2242–50. http://dx.doi.org/10.1109/TIE.2015.2510505.

[36] Du X, Htet KKK, Tan KK. Development of a genetic-algorithm-based nonlinear model predictive control scheme on velocity and steering of autonomous vehicles. IEEE Trans Ind Electron 2016;63(11):6970–7. http://dx.doi.org/10.1109/TIE.2016.2585079.

[37] Halim M, Khaled B. Accelerated micro particle swarm optimization for the solution of nonlinear model predictive control. World J Eng 2017;14(6):509–21. http://dx.doi.org/10.1108/WJE-01-2017-0004.

[38] Li Z, Deng J, Lu R, Xu Y, Bai J, Su CY. Trajectory-tracking control of mobile robot systems incorporating neural-dynamic optimized model predictive approach. IEEE Trans Syst, Man Cybern:Syst 2016;46(6):740–9. http://dx.doi.org/10.1109/TSMC.2015.2465352.

[39] Li Z, Xia Y, Su C-Y, Deng J, Fu J, He W. Missile guidance law based on robust model predictive control using neural-network optimization. IEEE Trans Neural Netw Learn Syst 2015;26(8):1803–9. http://dx.doi.org/10.1109/TNNLS.2014.2345734.

[40] Jacobson B. On the mean value theorem for integrals. Amer Math Monthly 1982;89(5):300–1. http://dx.doi.org/10.1080/00029890.1982.11995437.

[41] He W, Dong Y, Sun C. Adaptive neural impedance control of a robotic manipulator with input saturation. IEEE Trans Syst Man Cybern: Syst 2016;46(3):334–44. http://dx.doi.org/10.1109/TSMC.2015.2429555.

[42] Ge SS, Wang C. Adaptive neural control of uncertain MIMO nonlinear systems. IEEE Trans Neural Netw 2004;15(3):674–92. http://dx.doi.org/10.1109/TNN.2004.826130.

[43] Modares H, Lewis FL. Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. Automatica 2014;50(7):1780–92. http://dx.doi.org/10.1016/j.automatica.2014.05.011.

[44] Kyriakos, Vamvoudakis G. Event-triggered optimal adaptive control algorithm for continuous-time nonlinear systems. IEEE/CAA J Autom Sin 2014;1(3):282–93. http://dx.doi.org/10.1109/JAS.2014.7004686.

[45] Wang D, Mu C. Adaptive-critic-based robust trajectory tracking of uncertain dynamics and its application to a spring-mass-damper system. IEEE Trans Ind Electron 2018;65(1):654–63. http://dx.doi.org/10.1109/TIE.2017.2722424.