Article

Constrained DNN-Based Robust Model Predictive Control Scheme with Adjustable Error Tube

Shizhong Yang, Yanli Liu * and Huidong Cao

College of Information and Control Engineering, Qingdao University of Technology, Qingdao 266525, China
* Correspondence: grace_liu345@sina.com

Abstract: This paper proposes a novel robust model predictive control (RMPC) scheme for constrained linear discrete-time systems with bounded disturbance. Firstly, the adjustable error tube set, which is affected by local error and error variety rate, is introduced to overcome uncertainties and disturbances. Secondly, the auxiliary control rate associated with the cost function is designed to minimize the discrepancy between the actual system and the nominal system. Finally, a constrained deep neural network (DNN) architecture with symmetry properties is developed to address the optimal control problem (OCP) within the constrained system while conducting a thorough convergence analysis. These innovations enable more flexible adjustments of state and control tube cross-sections and significantly improve optimization speed compared to the homothetic tube MPC. Moreover, the effectiveness and practicability of the proposed optimal control strategy are illustrated by two numerical simulations. In practical terms, for 2-D systems, this approach achieves a remarkable 726.23-fold improvement in optimization speed, and for 4-D problems, it demonstrates an even more impressive 7218.07-fold enhancement.

Keywords: symmetry; constrained system; robust model predictive control; deep neural network; bounded disturbance

1. Introduction

Over the past few years, RMPC was enjoying enormous acceptance in practical applications, including trajectory tracking, industrial process control, and energy systems [1–3]. The successful implementation of RMPC in the various branches is on account of its prominent advantages. In particular, RMPC provides an integrated solution for controlling systems with model uncertainty, additive disturbance, and constraints. Theoretically, the feature attracted remarkable attention for analyzing and synthesizing different forms of RMPC. As a result, several RMPC algorithms were investigated in the literature [4–7], and so on.

Recently, the application requirements for considering practical constraints and the realization environment prompted increasing attention of RMPC towards new orientations. For instance, the increasing demand for algorithms underscores the need to integrate optimization performance and control robustness, propelling the development of tube-based MPC (TMPC) [8–10]. The deployment of tubes draws forth a set of strictly set theoretic strategies for RMPC synthesis, which consider a computationally efficient treatment of uncertainties and their interaction with the system dynamics and constraints. In [11,12], a class of linear systems with bounded disturbance and convex constraint separated the nominal system from the actual system by adopting a separation control strategy. What is noteworthy is that the conservatism of the proposal employing this construction in [13,14] was caused by deploying the fixed tube cross-section shape sets. To mitigate this conservatism, the homothetic tube model predictive control (HTMPC) strategy proposed in [15,16] explored the impact of disturbances by constructing locally accurate reachable sets centered around nominal system trajectories. In light of these developments, the concept of HTMPC

emerged as an enhanced and more adaptable framework for RMPC synthesis. Among the array of control schemes considered, HTMPC stands out as an improved and more versatile option. What sets it apart is its capacity to parameterize the cross-sections of the state tube and control tube in terms of associated centers and scaling sequences. This study aims to further investigate this concept by considering variations in state error and changes in the value of the cost function during error adjustment in designing the tube size controller and auxiliary control law, which distinguishes it from previous literature [15] that incorporates scaling vector optimization into OCP, thereby increasing computational complexity and aiming to optimize scaling vectors to a specific value. However, it is essential to note that an inherent drawback of the HTMPC approach lies in its computational complexity, which grows significantly with an increasing number of constraints, as measured by the proliferation of polytopic regions.

Furthermore, the issue of computational complexity is generally associated with dynamic programming in the presence of constraints and uncertainties, which inspires the development of parameterized RMPC [17–19]. The parameterized optimization problem is commonly approximated using neural network (NN) or DNN to enhance computational efficiency [20,21]. Certain studies even turned to symmetric neural networks (SNNs) due to their unique properties [22]. SNNs, characterized by symmetric weight initialization and activation functions, demonstrated their ability to accelerate convergence and improve the robustness of neural network-based approaches [23]. Some studies adopted an offline approach to generate nominal systems [24,25]. While effective in reducing online computation time, this method leans toward a more conservative control strategy in highly uncertain scenarios, necessitating a trade-off with control performance. Additionally, other studies considered system uncertainty by establishing a linear variable parameter system model [26,27]. This approach facilitates adaptive learning to address system changes and uncertainties, making it better suited for handling variations in the variable parameter system. However, applying this technique in complex, large-scale systems demands substantial computational resources for the training and inference of DNN, potentially leading to real-time control delays. As the field of online learning technology continues to mature, its integration with RMPC holds promise for enhancing the real-time capabilities and scalability of the control scheme. Notably, previous studies employed reinforcement learning techniques to solve linear quadratic regulator and MPC problems, providing convergence proofs for associated issues [28,29]. Advanced deep reinforcement learning algorithms further demonstrated their potential within an RMPC framework, emphasizing the iterative interaction between optimal control actions and performance indices [30,31]. These instances underscore the capacity of online learning techniques to address quadratic programming problems. Therefore, the integration of online learning techniques, including deep neural networks (DNNs) with a symmetric architecture, holds immense potential in enhancing the real-time capabilities and scalability of robust model predictive control (RMPC). Our proposed approach, which leverages the computational power of GPUs for real-time acquisition of time-varying nominal system information, not only ensures real-time control performance, but also optimizes efficiency.

Building upon the above research, it is not difficult to find that a promising approach involves incorporating tubes with increased degrees of freedom into the optimization process while employing function approximation and online learning techniques within the framework of RMPC to enhance computational efficiency. The main contributions of the paper are three-fold:

1.　A fuzzy-based tube size controller is investigated to adjust the local error tube-scaling vector. Specifically, the controller is designed by considering the state error between the nominal and the actual systems; the error and error variety rate bounds are then established, and the fuzzy IF-THEN rules are derived. The tightened sets on state error are developed to satisfy the system constraints in the case of external disturbances and model uncertainties.

2. An auxiliary control law pertaining to the scaling vector of the error tube holds greater significance. The auxiliary control law effectively mitigates interference impact on the system by considering variations in the system's cost function.

3. A theoretically rigorous and technically achievable framework for RMPC with online parameter estimation, based on a constrained DNN with symmetry properties to improve computing performance, was developed: the OPC is defined based on the parameters of online learning; the DNN structure is expanded using Dykstra's projection algorithm to ensure the feasibility of the successor state and control input; a time-varying nominal system is generated based on the aforementioned content to fulfill the requirements of system robustness.

The remainder of this paper is organized as follows: Preliminaries and problem formulation is considered in Section 2. In Section 3, a novel RMPC scheme is developed based on the fuzzy-based tube size controller and constrained DNN algorithm. Section 4 provides two numerical examples to illustrate the feasibility and effectiveness of the proposed control scheme. In Section 5, some conclusions are drawn.

2. Preliminaries and Problem Formulation

2.1. Nomenclatures

The set of non-negative reals is denoted by $\mathbb{R}$; ℓ_N is a sequence of non-negative integers $\ell_N \triangleq \{0, 1, 2, \ldots, N\}$. For a set A and a real matrix M of compatible dimensions, the image of A under M is denoted by $MA = \{Ma : a \in A\}$. Given two subsets C and B of $\mathbb{R}^n$ and $x \in \mathbb{R}^n$, the Minkowski set addition is defined by $C \oplus B \triangleq \{c + b | c \in C, b \in B\}$ and Minkowski set subtraction is defined by $C \ominus B \triangleq \{c | c \oplus B \subseteq C\}$. $\{x\} \oplus C$ is substituted for $x \oplus C$. For $M > 0$ and $x \in \mathbb{R}^n$, define $\|x\|_M^2 = x^T M x$. The distance of a point $x \in \mathbb{R}^n$ from a point $z \subset \mathbb{R}^n$ is denoted by $d(x, z) = |x - z|$. $Conv\{\cdot\}$ denotes the convex hall of elements in $\{\cdot\}$. For an unknown vector v, the notations v^* represent its optimal value.

2.2. Problem Formulation

Consider a discrete-time linear system with bounded disturbance (actual system) in the form of

$$x_{k+1} = Ax_k + Bu_k + w_k, \ k \in \ell_N, \tag{1}$$

where N is the horizon length. $x_k \in \mathbb{R}^n$ and $u_k \in \mathbb{R}^m$ are the state vector and the control input of the actual system subject to bounded disturbance w_k. $w_k \in \mathbb{R}^n$ is taking values in the set $\mathbb{W} \in \mathbb{R}^n$. The x_{k+1} denotes the successor state of the actual system. The system variables are selected such that the following constraints are satisfied:

$$x_k \in \mathbb{X} \subseteq \mathbb{R}^n, u_k \in \mathbb{U} \subseteq \mathbb{R}^m, w_k \in \mathbb{W} \subseteq \mathbb{R}^n, \ k \in \ell_N, \tag{2}$$

where $\mathbb{X}$ and $\mathbb{U}$ are compact and convex, which contains the origin as an interior point. The compact set $\mathbb{W}$ contains the origin.

Let the nominal (reference) system without any disturbance corresponding to (1) be defined by

$$z_{k+1} = Az_k + Bv_k, \ k \in \ell_N, \tag{3}$$

where $z_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ are the state and control input of the nominal system without accounting for any uncertainty, respectively. z_{k+1} denotes the desired value of the successor state in the system (1).

The state error is represented as

$$e_k = x_k - z_k, \ k \in \ell_N. \tag{4}$$

Assumption 1.

- *The matrix pair $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$ is known and stabilizable;*
- *The state x_k can be measured at each sample time;*

- *The current disturbance $w_k \in \mathbb{W}$ and future disturbances $w_{k+i} \in \mathbb{W}, i = 1, 2, \cdots, N - 1$ are not known and can take arbitrary values.*

In this paper, the fixed shape set of the error tube is expressed as E. For any non-empty set $E \subseteq \mathbb{R}^n$, the error tube is a sequence of sets $\mathbf{E}_N = \{E_k\}$, where E_k is given by

$$E_k = \alpha_k E, \ k \in \ell_N \text{with} \alpha_k \in \mathbb{R}, \tag{5}$$

where α_k is the scaling vector. Meanwhile, for each relevant $k \in \ell_N$, the state tube $\mathbf{X}_N$ and control tube $\mathbf{U}_{N-1}$ corresponding to HTMPC [18] are indirectly determined by the following form

$$\mathbf{X}_N = \{z_k(e_{k-1})\} \oplus \mathbf{E}_N, \ k \in \ell_N \tag{6}$$

$$\mathbf{U}_{N-1} = \{v_k(e_k)\} \oplus K\mathbf{E}_N, \ k \in \ell_N \tag{7}$$

where $\{z_k(e_{k-1})\}$ and $\{v_k(e_k)\}$ are the sequence of state tube and control tube centers determined by state error e. $K \in \mathbb{R}^{m \times n}$ is the disturbance rejection gain [32]. The corresponding control policy is a sequence of control laws $\mathbf{\Pi}_{N-1} = \{\pi_k(e_k, E_k, U_k)\}$ with

$$\forall e_k \in \alpha_k E, \pi_k(e_k, E_k, U_k) = v_k(e) + Ke_k, k \in \ell_{N-1}. \tag{8}$$

Refer to Equations (5)–(8), clearly, given set E, the error tube $\mathbf{E}_N$, state tube $\mathbf{X}_N$, control tube $\mathbf{U}_{N-1}$, and control policy $\mathbf{\Pi}_{N-1}$ are determined by the sequences of $\{e_k \in \mathbb{R}^n\}$ and $\{v_k \in \mathbb{R}^m\}$. Consequently, introduce a decision variable $\varphi_N = (e_0, \ldots, e_N, v_0, \ldots, v_{N-1}) \in \mathbb{R}^{N(n+m+1)}$.

Subsequently, the OCP $\mathbb{P}_N(e)$ is defined by

$$V_N^0(e) = \inf_{\varphi_N} \{V_N(\varphi_N) : \varphi_N \in \Phi_N(e)\}, \tag{9}$$

$$d_N^0(e) = \operatorname*{arginf}_{\varphi_N} \{V_N(\varphi_N) : \varphi_N \in \Phi_N(e)\}, \tag{10}$$

where the cost function $V_N(\cdot)$ is defined by

$$V_N(\varphi_N) = \min \sum_{k=0}^{N-1} \ell(e_k, v_k) + V_f(e_N), \tag{11}$$

with

$$\ell(e_k, v_k) = \|e\|_{Q_e}^2 + \|v\|_{Q_v}^2, k \in \ell_N \tag{12}$$

and

$$V_f(e_N) = \|e\|_P^2, \tag{13}$$

here, $\ell(e_k, v_k)$ is the stage cost, which is employed to achieve the desired performance of the control. The terminal cost represented by $V_f(e_N)$ ensures stability and recursive feasibility. $Q_e \in \mathbb{R}^{n \times n}$, $Q_v \in \mathbb{R}^{m \times m}$, and $P \in \mathbb{R}^{n \times n}$ are known positive definite symmetric matrices. For any $x_k \in \mathbb{X} \subseteq \mathbb{R}^n$, the set of permissible decision variables φ_N corresponds to the value of the set-valued map $\Phi_N(e)$ as $\Phi_N(e) := \{\varphi_N : (14) \text{ holds for all } k \in \ell_{N-1}\}$, where

$$e_0 \in E, \tag{14}$$

$$(A + BK)e \oplus \mathbb{W} \subseteq \alpha_{k+1} E, \tag{15}$$

$$\forall e_k \in \alpha_k E, \tag{16}$$

$$\{\upsilon(e_k)\} \in \mathbb{U} \ominus \alpha_k KE, \tag{17}$$

$$\{z(e_{k-1})\} \in \mathbb{X} \ominus \alpha_k E, \tag{18}$$

$$Ax_k + Bu_k \oplus \mathbb{W} \subseteq z(e_k) \oplus \alpha_{k+1} E, \tag{19}$$

$$E_N \in E_f, \tag{20}$$

where $E_f \subseteq \mathbb{R}^{n+1}$ is the terminal constraint set [33] for $\mathbb{P}_N(e)$.

Similar to the tube MPC principle [8], if z_k satisfies $\mathbb{X} \ominus \alpha_{k+1}E$ and v_k satisfies $\mathbb{U} \ominus KE$, then the imposed constraints on the actual system state $x \in \mathbb{X}$ and control input $u \in \mathbb{U}$ are also met. In this work, the determination of z_k is related to e_{k-1}, while the determination of v_k is concerned with e_k; thus, it is imperative to satisfy both conditions $\{z(e_{k-1})\} \in \mathbb{X} \ominus \alpha_k E$ and $\{v(e_k)\} \in \mathbb{U} \ominus \alpha_k KE$. Furthermore, at step N, if E_N fulfills terminal constraint $E_N \in E_f$ (the Equation (21) provides the formulation and limitation of E_f), it guarantees that the system state complies with requirement $x_N \in \mathbb{X}$.

Constraints (15) and (19) represent the set dynamics of the error tube and the homothetic state tube, respectively, which contribute to dynamic relaxation in [8]. In addition, the terminal constraint set E_f satisfies the following constraint:

$$(A + BK)E_f \subset E_f. \tag{21}$$

The performance evaluation of the terminal control necessitates the definition of a 0-step homothetic tube controllability set X_0 [15], which must satisfy the following constraints:

$$X_0 = \mathrm{Proj}_{\mathbb{R}^n}\{(x, z, \alpha) : z \in \mathbb{X} \ominus \alpha E | Kz \in \mathbb{U} \ominus \alpha E\}. \tag{22}$$

where $\mathrm{Proj}_{\mathbb{R}^n}(Z)$ denotes a set $z \subset \mathbb{R}^{n+m}$ projected onto $\mathbb{R}^n$ as $\mathrm{Proj}_{\mathbb{R}^n}(Z) = \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^m such that (x, y) \in Z\}$.

2.3. Controller Synthesis

The objectives of this paper is to design an optimal control policy u_k based on any given initial state error e_0, which not only renders the local state x_k asymptotically tracking the reference state z_k, namely e_k asymptotically approaching zero, but also minimizes the OCP. The problem of solving the conventional control policy u_k of (1) is converted into finding the nominal optimal control input $v_k(e)$ and designing an appropriate disturbance rejection gain K while ensuring that the constraints related to α_k are satisfied.

The controller synthesis for the proposed RMPC scheme is specified as

$$u_k = v(e_k) + Ke_k, \ k \in \ell_N \tag{23}$$

where u_k is the control action obtained from the presented method. The ancillary control law is denoted as Ke_k, which keeps the local state x_k within the error tube centered around the trajectory of z_k. $v(e_k)$ is the output obtained by online learning with state errors as input.

Consider the error system obtained by integrating the Equations (1), (3), and (4) as

$$e_{k+1} = Ae_k + B(u_k - v_k) + w_k, \ k \in \ell_N \tag{24}$$

where e_{k+1} is the successor state error. The system (24) is rewritten to be

$$e_{k+1} = (A + BK)e_k + w_k, \ k \in \ell_N. \tag{25}$$

3. DNN-Based RMPC with a Fuzzy-Based Tube Size Controller

This section presents the design of the novel RMPC scheme, which incorporates updates to scaling and policy iteration for nominal control. The innovative RMPC framework consists of a fuzzy-based tube size controller and a constrained DNN-based nominal RMPC component. The former calculates the error tube-scaling vector by considering both state

error and error variety rate, while the latter determines a sequence of constraints associated with scaling to ensure optimal control policy generation. Concurrently, the DNN-based nominal RMPC offers a time-varying nominal system that exhibits enhanced computational efficiency. Moreover, by incorporating variations in the cost function value into the auxiliary control law design, it effectively mitigates the adverse effects of interference on the system.

3.1. Error Tube and Constraint Satisfaction

This work discusses that the fuzzy control is used to estimate (predict) the corresponding error tube-scaling vector α_k, allowing for computational feasibility of the OCP $\mathbb{P}_N(e)$. More importantly, an auxiliary control law Ke_k pertaining to the scaling vector of the error tube holds greater significance. The auxiliary control law effectively mitigates interference impact on the system by considering variations in the system's cost function.

Assumption 2.

- *The error tube cross-section shape set $E \subset \mathbb{R}^n$ (i.e., outer invariant approximation of the minimal robust positively invariant set [34]) is compact, convex, and contains the origin such that $\{(A + BK)e : e \in E\} \oplus \mathbb{W} \subseteq \alpha_{k+1}E, k \in \ell_{N-1}$;*
- *The state tube cross-section shape set $\mathbb{Z}$ is given by $\mathbb{Z} = Conv\{\mathbb{Z}(e) : e \in E\}$;*
- *The control tube cross-section shape set $\mathbb{V}$ is given by $\mathbb{V} = Conv\{v(e) : e \in E\}$.*

If E satisfies Assumption 2, then for any established $\alpha_k \in \mathbb{R}$, it holds that $e_k \in \alpha_k E$. Further, the nominal state and control input are restrained indirectly as $\mathbb{Z}(e) \in \mathbb{Z}$ and $v(e) \in \mathbb{V}$. It is clear that if $\forall e_k \in \alpha_k E$, $v(e) \in \mathbb{V}$, then the satisfaction of original constraints $u \in \mathbb{U}$ for $\forall w \in \mathbb{W}$ is guaranteed by using the control scheme $u_k(e) = v(e_k) + Ke_k$.

Next, the fuzzy-based tube size controller is employed to estimate the error tube scaling, which generates the scaling vector by considering the local error and the error variety rate. The components of the fuzzy controller [35] include some fuzzy IF-THEN rules and a fuzzy inference engine. The fuzzy inference engine utilizes the IF-THEN rules to map from input error $e \in \mathbb{R}^n$ and error variety rate $e_c \in \mathbb{R}^n$ to an output variable $\alpha \in \mathbb{R}$. The lower and upper bound values of e and e_c are represented as $\pm e_0$ and $\pm e_0'$, respectively. Furthermore, divide the two-dimensional graph comprising e and e_c into nine distinct regions, as depicted in Figure 1. Upper and lower limits for both e and e_c define these regions. Each region, denoted as G_i, corresponds to a specific IF-THEN rule. The fuzzy controller accurately determines the region within the graph where a given pair of values for e and e_c are located, based on the provided input. Subsequently, it employs IF-THEN rules to calculate the appropriate scaling variables. Taking G_{3+} as an illustrative example, in this particular scenario, when $e \geq e_0$ and $e \geq e_0$, it indicates a relatively high positive deviation of the system's state error with a gradual increase. In such circumstances, the controller generates a diminished value for α as an output, ensuring that the system's state error exhibits a tightening trend.

To be specific, fuzzy IF-THEN rules are written as

1. (G_1). IF $e \leq -e_0$ and $e_c \geq e_0'$ or $e \geq e_0$ and $e_c \leq -e_0'$ THEN α takes on a smaller value;
2. (G_2). IF $|e| \leq e_0$ and $e_c \geq e_0'$ or $|e| \leq e_0$ and $e_c \leq -e_0'$ THEN takes on a slightly larger value;
3. (G_3). IF $e \leq -e_0$ and $e_c \leq |e_0'|$ or $e \geq e_0$ and $e_c \leq |e_0'|$ THEN α takes a value as small as possible.
4. (G_4). IF $|e_c| \leq e_0'$ and $e \geq e_0$ or $|e_c| \leq e_0'$ and $e \leq -e_0$ THEN α takes on a larger value;
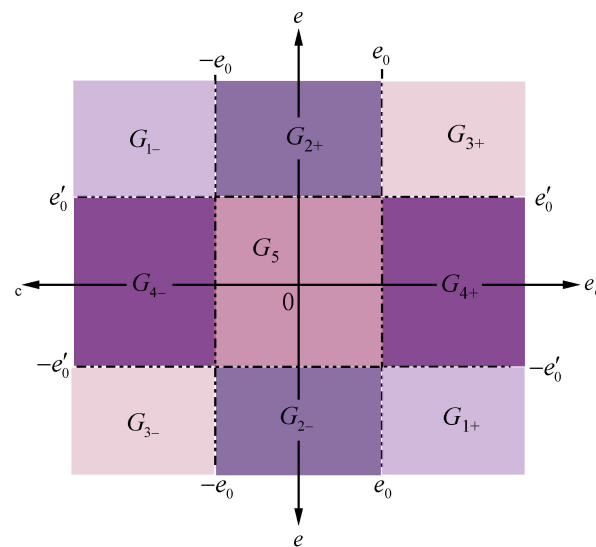5. (G_5). IF $|e| < e_0$ and $|e_c| < e_0'$ THEN α takes a value as large as possible.

Figure 1. A two-dimensional depiction of the control maneuver. Different regions in the graph represent the following: G_{1-} corresponds to the region where $e \leq -e_0$ and $e_c \geq e_0'$; G_{1+} depicts the area with $e \geq e_0$ and $e_c \leq -e_0'$; G_{2+} characterizes the area where $|e| \leq e_0$ and $e_c \geq e_0'$; G_{2-} highlights the territory where $|e| \leq e_0$ and $e_c \leq -e_0'$; G_{3+} showcases the domain where $e \geq e_0$ and $e_c \leq |e_0'|$; G_{3-} marks the domain where $e \leq -e_0$ and $e_c \leq |e_0'|$; G_{4+} exemplifies the region where $|e_c| \leq e_0'$ and $e \geq e_0$; G_{4-} describes the space where $|e_c| \leq e_0'$ and $e \leq -e_0$; G_5 specifies the domain where $|e| < e_0$ and $|e_c| < e_0'$.

For convenience, let the universe of e be a~b and set the universe of e_c as c~d. The membership degree function is taken as the triangular function. Then, singleton fuzzifier and average center defuzzifier [36] were used to calculate outputs α based on the feedback values of e and e_c in the form of

$$\alpha = \frac{\sum\limits_{i=1}^{5} \eta_i \mu(e, e_c)}{\sum\limits_{i=1}^{5} \mu(e, e_c)}, \tag{26}$$

where $\mu(\cdot)$ is the membership degree of the five cases mentioned above. The η_i is an adjustable weight parameter of α under a different context. Afterward, the successor value of α is determined by

$$\alpha_{k+1} = \alpha_k + \tau, \tag{27}$$

with

$$\tau = \max_{\lambda}\{\lambda | \mathbb{W} \in \lambda E\}. \tag{28}$$

Theorem 1. *Given system (1) controlled with the control policy $u \triangleq \{u_0, u_1, \ldots, u_{N-1}\}$, the state error e_k is restricted to the error tube $\alpha_k E$. To be specific, the design of the disturbance rejection rate ensures that error $\lim\limits_{k \to \infty} e_k \to 0$ for $\forall w \in \mathbb{W}$.*

Lemma 1 ([37]).

$$s_1{}^T F s_2 + s_2{}^T F s_1 \leq s_1{}^T F s_1 + s_2{}^T F s_2,$$

where s_1, s_2 are any vector. $F \in \mathbb{R}^{m \times n}$ is a positive definite matrix.

Proof of Theorem 1. Consider the error system (25). The disturbance rejection gain K guarantees that e_k is constrained to be inside the set $\alpha_k E$, i.e., $x_k \in z_k + \alpha_k E$. Since the nominal system (3) has robust stability, the nominal state z_k should converge to the

origin $d(z_k, 0) \to 0$. Then, the state error e_k must converge to error tube $\alpha_k E$ because of $x_k \in z(e_{k-1}) + \alpha_k E$, namely $d(e_k, \alpha_k E) \to 0$. Finally, the state error e_k is restricted to a variable error tube $\alpha_k E$ whose center is at the origin by implementing the ancillary control law Ke_k.

Here, the disturbance rejection gain K is solved by the following equation

$$K^T B^T H I_n + H I_n BK + 2\|BK\|^2_{HI_n} + \left(3\tau^2 + 1\right) \cdot H I_n = -Q \tag{29}$$

where H is determined by equation $H = (V_N(d_N))^T V_N(d_N)$ and Q is a positive definite matrix. I_n denotes the identity matrix with the same dimensions as the state vector x_k. For convenience, let us set $H I_n = P_V$.

Then, P is the solution to the following Lyaponuv equation

$$\varepsilon^2 \|A + BK\|^2_P - P = -Q_e - \|K\|^2_{Q_v} \tag{30}$$

with

$$\varepsilon \in \left(1, \frac{1}{Eig_{\max}(A + BK)}\right) \tag{31}$$

where $Eig_{\max}(\cdot)$ is the maximum value of the matrix eigenvalue.

The Lyapunov candidate function is represented as

$$V_L = e_k{}^T P_V e_k, \ k \in \ell_N. \tag{32}$$

Consider the first difference equation as

$$\Delta V_L = e_{k+1}{}^T P_V e_{k+1} - e_k{}^T P_V e_k. \tag{33}$$

By substituting Equation (25) into Equation (33), one obtains

$$\Delta V_L = (BKe_k)^T P_V e_k + (BKe_k)^T P_V BKe_k + (BKe_k)^T P_V w_k + w_k{}^T P_V e_k + w_k{}^T P_V BKe_k + w_k{}^T P_V w_k + e_k{}^T P_V BKe_k + e_k{}^T P_V w_k \tag{34}$$

According to Lemma 1, then it follows that

$$\Delta V_L \leq (BKe_k)^T P_V e_k + e_k{}^T P_V e_k + e_k{}^T P_V BKe_k + 2(BKe_k)^T P_V BKe_k + 3w_k{}^T P_V w_k. \tag{35}$$

Or, equivalently

$$\Delta V_L \leq e_k{}^T \left(K^T B^T P_V + 2\|BK\|^2_{P_V}\right) e_k + e_k{}^T P_V e_k + e_k{}^T P_V BKe_k + 3w_k{}^T P_V w_k. \tag{36}$$

According to Equation (28), the disturbance is bounded by τ as $w_k \leq \tau e_k$. We have

$$\Delta V_L \leq e_k{}^T \left(K^T B^T P_V + 2\|BK\|^2_{HP_V} + P_V BK + \left(3\tau^2 + 1\right) \cdot P_V\right) e_k. \tag{37}$$

By substituting Equation (29) into Inequation (37), further obtain

$$\Delta V_L \leq e_k{}^T (-Q) e_k. \tag{38}$$

It is clear that $\Delta V_L \leq 0$, thus the function (32) is a decreasing function, then $\lim_{t \to \infty} e_k \to 0$.

$\square$

This section shows that optimal cross-sections of the error tube are calculated online by considering the adjustable tube-scaling parameters α_k, which are affected by a combination of error and error variety ratio. Theorem 1 shows that the successor estimation of the actual system has a non-increasing estimation error at each time step. The design of the

fuzzy-based tube size controller and the auxiliary control law considers both variations in state error and changes in the value of cost function during error adjustment, unlike previous literature [15] that incorporates optimization of scaling vector α_k into OCP, thereby increasing computational complexity and aiming to optimize scaling vectors to a specific value. In addition, we discover that the appropriate selection of the acquisition form of a nominal system can improve prediction accuracy. Nonetheless, the invariable nominal system is considered during the prediction in [24,25]. In order to improve the control performance, our main concern here is to define a parameter estimation scheme that generates a time-varying nominal system based on the DNN algorithm and still enables a computationally tractable RMPC algorithm, which is presented in the following.

3.2. Design of DNN-Based Nominal RMPC

This section focuses on designing the DNN-based nominal RMPC to construct a parameter estimation synthesis that provides a time-varying nominal system for the control scheme. The cost function for the constrained system proposed in conventional RMPC is reformulated as an online learning problem by introducing a series of reference control inputs $v_k = v^\theta(e_k)$ parameterized by θ. The modified OCP $\mathbb{P}_N^\theta(e)$, solved online, is defined by

$$V_N^\theta(e) = \inf_{\varphi_N} \sum_{k=0}^{N-1} \left(\|e_k\|_{Q_e}^2 + \|v^\theta(e_k)\|_{Q_v}^2 \right) + \|(e_N)\|_P^2, \tag{39}$$

$$\varphi_N^\theta(e) = \operatorname*{arginf}_{\varphi_N} \sum_{k=0}^{N-1} \left(\|e_k\|_{Q_e}^2 + \|v^\theta(e_k)\|_{Q_v}^2 \right) + \|(e_N)\|_P^2. \tag{40}$$

$$\text{s.t. } \forall e_k \in \alpha_k E, \ k \in \ell_{N-1}, \tag{41}$$

$$(A + BK)e_k \oplus \mathbb{W} \subseteq \alpha_{k+1} E, \ k \in \ell_{N-1}, \tag{42}$$

$$\left\{ v^\theta(e_k) \right\} \in \mathbb{U} \ominus \alpha_k K E, \ k \in \ell_{N-1}, \tag{43}$$

$$\{z(e_k)\} \in \mathbb{X} \ominus \alpha_{k+1} E, \ k \in \ell_{N-1}, \tag{44}$$

$$Ax_k + Bu_k \oplus \mathbb{W} \subseteq z(e_k) \oplus \alpha_{k+1} E, \ k \in \ell_{N-1}, \tag{45}$$

$$E_N \in E_f. \tag{46}$$

The parameters θ will update in the direction of the gradient $\nabla_\theta V_N^\theta(e)$ of the cost function by adopting the policy gradient method. In the architecture of constrained DNN-based nominal RMPC, the state errors $(e_k, \dots, e_N)$ are used as input to create the optimal control policy $v^\theta(e_k)$ as the output of DNN.

This paper employs DNN characterized by inherent symmetry, which features symmetric weights, facilitating efficient parameter sharing. Consequently, the network demands fewer computational resources than conventional network structures, rendering them advantageous in resource-constrained environments. Assuming the network has L hidden layers, the layers 1 and L each consist of i neurons. The architecture of a deep neural network is illustrated in Figure 2.
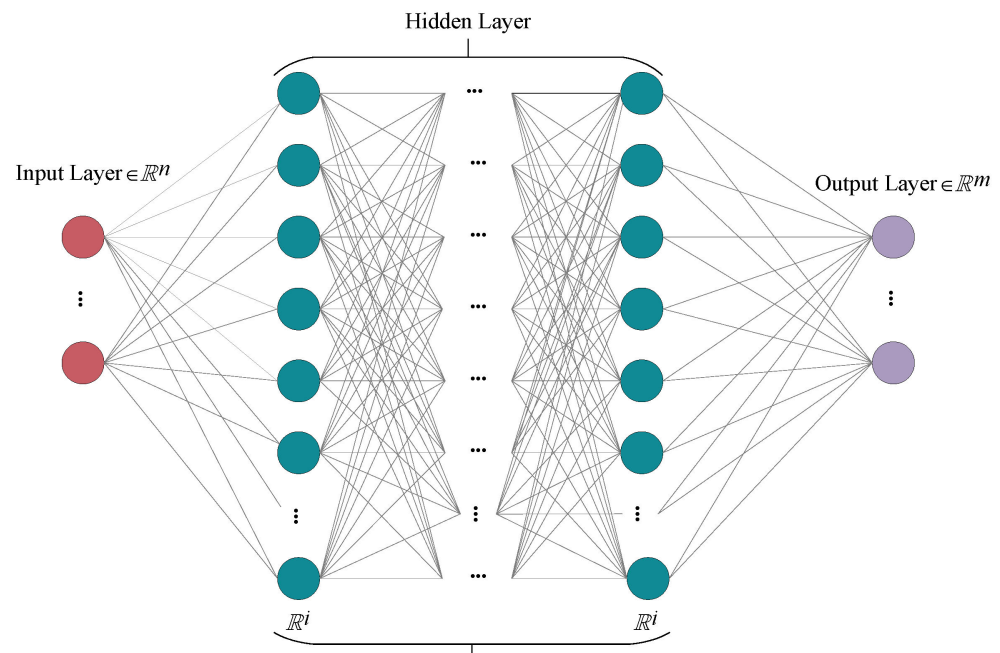
Figure 2. Diagram illustrating the architectural structure of the deep neural network.

The superiority of the network architecture employed in this paper over a typical neural network structure is demonstrated in Table 1.

Table 1. Performance comparison: symmetrical DNN vs. general DNN.

Network Structure	Number of Calculated Weights	Iterations	Precision
Symmetrical DNN	$2i(L+1)$	436	98.03%
Typical DNN	$[m+n+(i+1)L-i+1]i$	679	97.64%

From Table 1, it can be observed that in symmetric neural networks, the number of weights to be calculated is reduced since each connection is computed only once and then shared. Notably, despite having fewer parameters, deep neural networks with symmetric structures achieve higher accuracy under the same computational resources. Regarding convergence, symmetric neural networks require 35.79% fewer iterations than conventional neural networks.

The output of the DNN-based nominal RMPC is formulated as

$$v^{\theta}(e_k) = \delta\left(\sum_{i=1}^{m} W^L a^{L-1} + b^L\right), \tag{47}$$

where the linear relationship coefficient matrix and bias vector between the hidden layer and the output layer are denoted as $W \subset \mathbb{R}^{n \times m}$ and $b \subset \mathbb{R}^{n \times 1}$, respectively. The affine function parameters $\theta = \{W_{1:L}, b_{1:L}\}$ will be optimized. δ is a rectified linear unit function. The output value of the hidden layer is $a^L \subset \mathbb{R}^{m \times 1}$, and set the input a^1 to e_k.

Since the neural network may output a potentially infeasible $v^{\theta}(e_k)$ for a given error e_k, Dykstra's projection algorithm [38] is introduced to ensure that subsequent states and controls remain feasible. Its structure is shown in Figure 3.
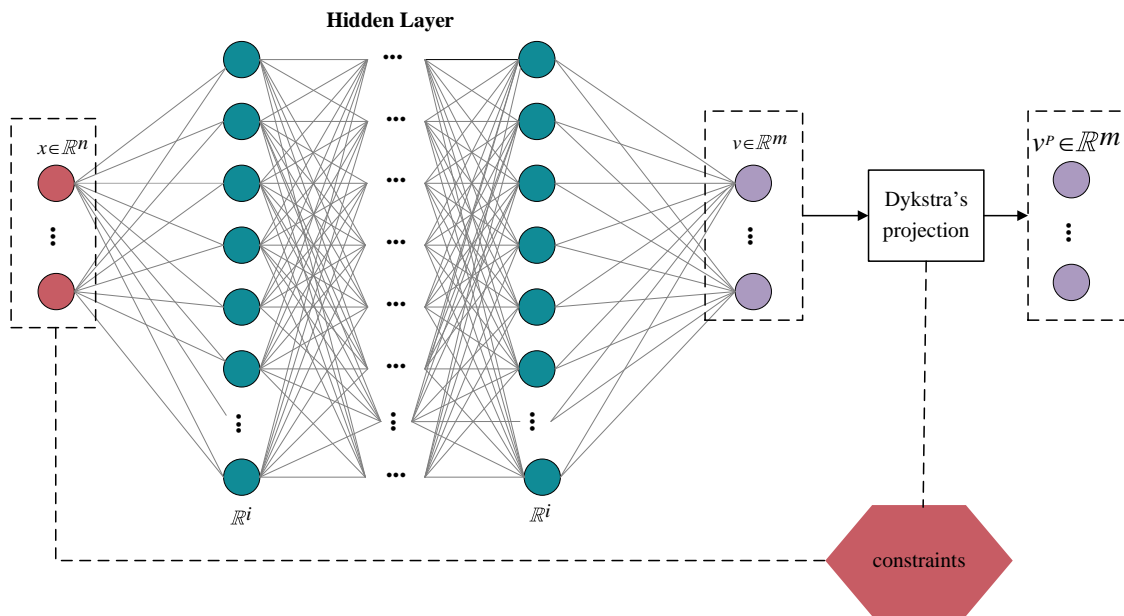
Figure 3. Diagram of the constrained DNN structure expanded by Algorithm 1.

Theorem 2. *By applying Dykstra's projection algorithm, the optimal control input $\upsilon^\theta(e_k)$ converges to the orthogonal projection of $\upsilon^\theta(e_k)$ onto the polytopic $\mathbb{U} \ominus \alpha_k KE$ as $t \to \infty$.*

Proof of Theorem 2. First, define the orthogonal projection of $\upsilon^\theta(e_k)$ onto the polytopic $\mathbb{U} \ominus \alpha_k KE$ as $P(\upsilon^\theta(e_k))$, and a series of variables $v^{(k,t)}$ and $I^{(k,t)}$ are generated from the DNN structure, which is extended by Dykstra's projection algorithm. It then iterates as

$$v^{(k,t)} = P\left(\upsilon^\theta(e_k)\right)\left(v^{(k-1,t)} - I^{(k,t-1)}\right), \tag{48}$$

$$I^{(k,t)} = v^{(k,t)} - \left(v^{(k-1,t)} - I^{(k,t-1)}\right). \tag{49}$$

Assume that the starting condition of the algorithm is $v^{(0,0)} = v(e_0)$ and $I^{(0,0)} = 0$. When $t \to \infty$, we have $I^{(k,t)} = I^{(k,t-1)}$, it is clearly that $v^{(k,t)} = v^{(k+1,t)}$ (i.e., the nominal control input $\upsilon^\theta(e_k)$ converges to $P(\upsilon^\theta(e_k))$). $\square$

Thus, given a state error e_k, control policy will output $P(\upsilon^\theta(e_k)) = f(e_k; \theta)$.

According to the policy gradient theory presented in [39], the gradient of the value function $\nabla_\theta V_N^\theta(e)$ with respect to the policy parameters θ is

$$\nabla_\theta V_N^\theta(e) = \mathbb{E}_v\left[V_N^\theta(e) \nabla_\theta \log \phi(v_t; f(e_t; \theta_t), \Sigma)\right], \tag{50}$$

where $\phi(v_t; f(e_t; \theta_t), \Sigma)$ is a multivariate Gaussian probability density function used to sample control inputs $\upsilon^\theta(e_k)$, centered at the DNN output $f(e_k; \theta)$ with diagonal covariance Σ, the covariance Σ anneals to 0 at the end of training to return to the control police.

The neural network parameters iterate by using stochastic gradient descent as

$$\theta_{t+1} = \theta_t - \gamma_t V_N^\theta(e) \nabla_{\theta_t} \log \phi(v_t; f(e_t; \theta_t), \Sigma). \tag{51}$$

The learning rate γ_t of DNN is selected as a positive number.

The termination criterion for the iteration is defined as

$$|\theta_{t+1} - \theta_t| \leq \left|V_N^\theta(e) - V_{N-1}^\theta(e)\right|. \tag{52}$$

For application of the proposed approach, instead of focusing on constructing a set of polytopic regions, function approximation and reinforcement learning techniques are used to directly learn an approximate optimal control policy. Furthermore, the policy gradient method guarantees the control action converges to locally optimal solutions by applying function approximation to generate unbiased estimates of the gradient with respect to the parameter θ. The proposed optimization method significantly enhances the computational performance of the system control while ensuring the feasibility of control inputs.

3.3. The Feedback Mechanism of the Control Synthesis

In this paper, the feedback loop encompasses state error, state error variety rate, and cost function, as illustrated in Figure 4. Expressly, the state error and error variety rate are conveyed to the fuzzy controller, subsequently yielding an error scaling vector associated with constraints at the subsequent time step. Simultaneously, the state error contributes to the optimization process of the cost function. The resulting cost function value is then fed back into the auxiliary control law, thereby determining the auxiliary control rate for the upcoming sampling time.
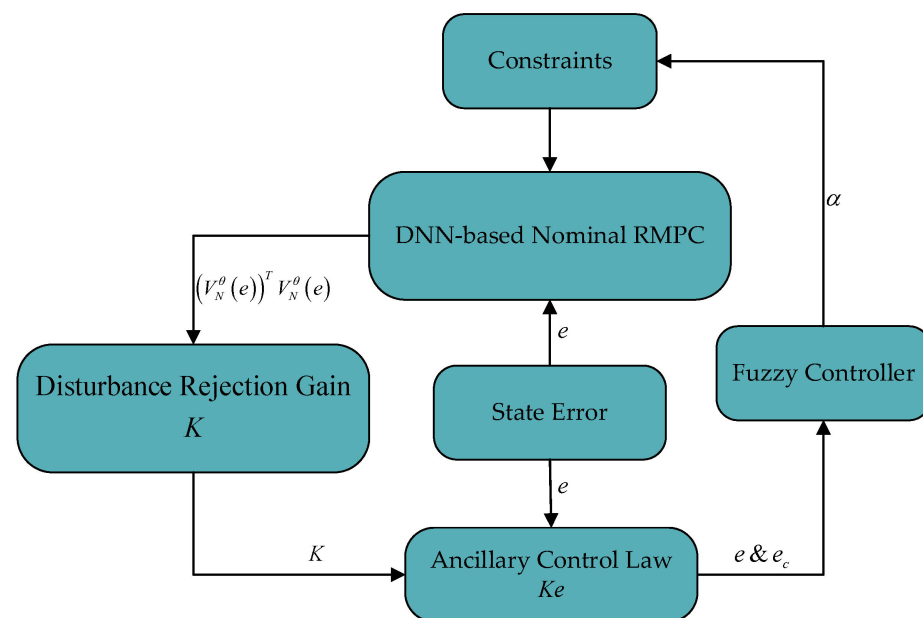

Figure 4. The flowchart depicting the feedback processing mechanism of the control synthesis.

The comparison between the computational performance of the proposed algorithm and HTMPC is shown in Table 1. Where $q_{\mathbb{X}}$, $q_{\mathbb{U}}$, q_E, and q_{E_f} in Table 1 denote the numbers of affine inequalities of the irreducible representation of the sets $\mathbb{X}$, $\mathbb{U}$, E, and E_f employed in the propose scheme; q_S and q_{G_f} are the numbers of affine inequalities of the irreducible representation of the state homothetic set and the terminal constraint set, respectively.

Table 2 clearly demonstrates that assigning the scaling vector to the fuzzy controller's specialized treatment not only provides a more comprehensive consideration of the impact of state error and error variety rate on error tube scaling, but also effectively reduces the number of decision variables and inequality constraints in the optimization process. Furthermore, the design of a symmetric constrained DNN structure addresses the issue of the exponential growth of polyhedra construction with the increasing number of constraints during the optimization. Consequently, implementing the proposed algorithm allows for a substantial reduction in computational complexity while enhancing the flexibility of system control.

Table 2. The comparison of computational complexity between the proposed approach and HTMPC.

Control Strategy	Number of Decision Variables	Number of Inequality Constraints	Upper Bound on the Number of Critical Regions
Proposed Approach	$N(m+n)+n$	$N(q_{\mathbb{X}}+q_{\mathbb{U}}+q_E)+q_{E_f}$	0
HTMPC	$N(m+n+1)+n+1$	$N(q_{\mathbb{X}}+q_{\mathbb{U}}+q_S+1)+q_S+q_{G_f}$	$2^{N(q_{\mathbb{X}}+q_{\mathbb{U}}+q_S+1)+q_S+q_{G_f}}$

3.4. The DNN-Based RMPC with a Fuzzy-Based Tube Size Controller Structure

To recapitulate, the proposed RMPC scheme comprises a fuzzy-based tube size controller and a DNN-based nominal RMPC part. The fuzzy-based tube size controller is employed to adjust error tube scaling. Meanwhile, the tightened sets (i.e., the minimal disturbance invariant set with an adjustable parameter α_k) and disturbance rejection gain K are computed online to restrain state error. Then, the DNN nominal RMPC is used to generate the time-varying nominal system in the case that the constraints associated with α are satisfied. It provides a theoretically rigorous and technically achievable framework for RMPC with online parameter estimation to improve calculated performance.

In this paper, we obtain the error tube shape set E by computing the minimum robust positively invariant set using the method described in [34]. Moreover, we set the error bound e_0 to 3.7 and secured the rate of error variation e_0' by 2.5. As parameter α typically ranges between 1 and 2, we design the fuzzy rule table in the Table 3 format. Table 3 shows that inputting a data pair (e_0, e_0') determines a reasonable value for α, which subsequently dictates α_{k+1}'s value according to Equation (27). The determination of α_{k+1} further influences determining associated constraints $\alpha_{k+1}E$, $\alpha_{k+1}KE$, $\mathbb{X} \ominus \alpha_{k+1}E$, and $\mathbb{U} \ominus \alpha_{k+1}KE$. Additionally, disturbance rejection rate and terminal cost function for control can be determined based on Equations (29) and (30). To meet specified constraints, constrained time-varying nominal system trajectories are computed through a Dykstra's projection algorithm-extended constrained DNN. The actual system will track the nominal trajectory while satisfying relevant conditions. Section 4 will explicitly discuss DNN parameter settings depending on the dimensionality of input and output variables. Specifically, Algorithm 1 gives the main procedure of the proposed control scheme, and its whole structure diagram is presented in Figure 5.

Algorithm 1 DNN-based RMPC with a fuzzy-based tube size controller

---

Given initial conditions $e_0 = 0$, $\alpha_0 = 1$ and weighting matrices Q_e, Q_v, determine the set E.
Compute the terminal weight matrix P and disturbance rejection gain K by using (29) and (30).
1: Randomly initialize θ
2: Set learning rate γ
3: **for** each time instant $k = 0,1,2,\ldots,$N **do**
4:     Compute polytopic $\alpha_k E$, $\alpha_k KE$, $\mathbb{X} \ominus \alpha_k E$ and $\mathbb{U} \ominus \alpha_k KE$
5:     **if** constraints (41)–(46) are satisfied **then**
6:         **repeat** calculate θ_{t+1} by using (51)
7:         **until** convergence
8:     **else**
9:         let $e_{k+1} = e_k$
10:     **end if**
11: Solve the optimization problem (39) and (40) based on θ_{t+1} to obtain $v_k^*(e_k)$,
12: Compute the error variety rate e_c and the corresponding scaling vector α, then obtain the successor scaling vector α_{t+1} by using (27),
13: Calculate the control input as $u_k(e) = \upsilon^\theta(e_k) + Ke_k$, and then implement u_k to the system.
14: **end for**

Table 3. Fuzzy rule comparison table.

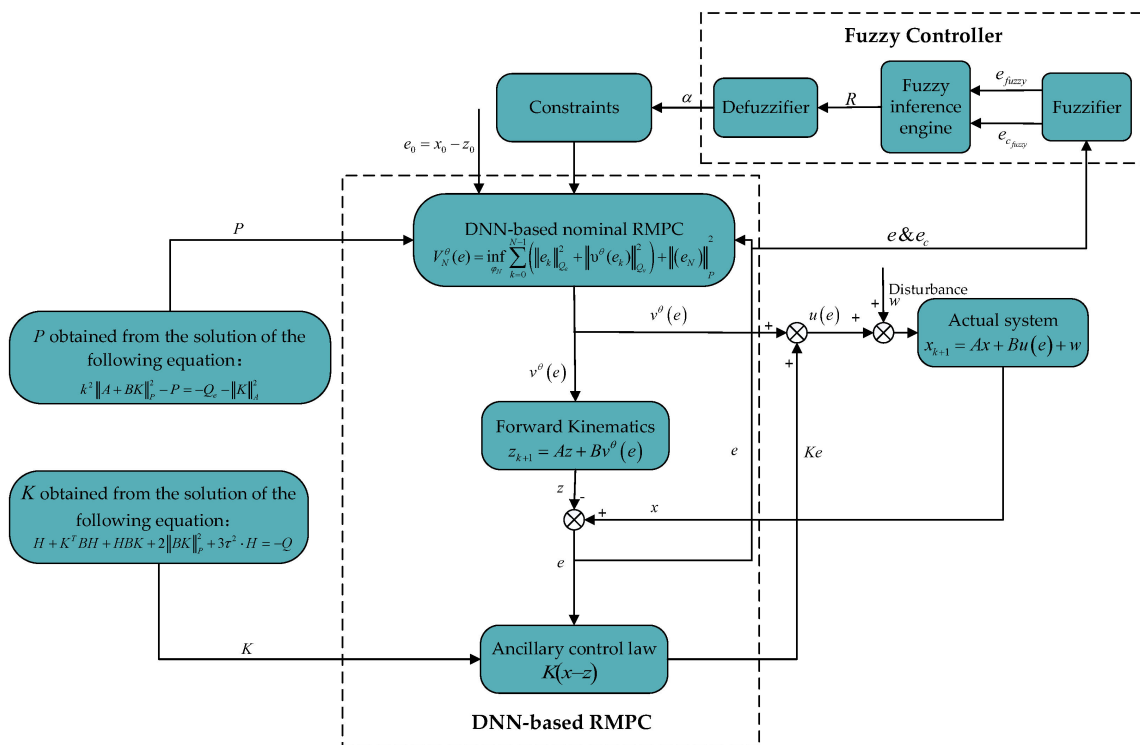| Scaling Vector	Fuzzy Rule Control Region				
	G_1	G_2	G_3	G_4	G_5
α	0.5–0.8	0.9–1.2	0–0.4	1.3–1.6	1.7–2.0

Figure 5. The structure of the constrained DNN-based robust model predictive control scheme with an adjustable error tube.

4. Simulations and Comparison Study

In this section, the advantages of the Algorithm 1 are illustrated by the following simulation examples of both 2-D and 4-D systems. The simulation experiments were conducted using Matlab, and the polyhedral constraint set was constructed utilizing the Mosek and MPT toolbox. Subsequently, the convex optimization problem of the actual system was solved. Deep learning toolboxes were employed to train neural networks for determining optimal control inputs in a nominal system.

Example 1. *Consider a 2-D double integrator discrete-time system in the form of (1) with*

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \ B = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}. \tag{53}$$

The state constraints are $x \in \mathbb{X} \triangleq \{[10, -2] \times [-10, 2]\}$, the disturbance is bounded as $w \in \mathbb{W} \triangleq \{w | \|w\|_\infty \leq 0.1\}$, and the control constraint is $u \in \mathbb{U} \triangleq \{u | |u| \leq 1\}$. The performance index function is defined in (39)–(46) with $Q_e = I_2$ and $Q_v = 0.01$, the terminal cost $V_f(e)$ is the value function $\|e_N\|_P^2$, while P is calculated from (30). Then, disturbance rejection gain K is computed by using (29). The set E is computed as a polytopic. The horizon length is selected as $N = 12$. The system is simulated using the initial condition $x_0 = z_0 = (-4, -2)$ and $\alpha_0 = 1$, the value of α_{k+1} is induced by Equation (27). In the context of neural network architecture determination, Figure 6 compares system nominal state trajectories when employing different network structures and deep neural networks with varying layers.

Indeed, from Figure 6, it is evident that when utilizing a symmetric DNN with six hidden layers, the trajectories of system nominal states can reach the desired values more rapidly (i.e., the trajectories of system nominal states reaching the origin by the 12th sampling time).
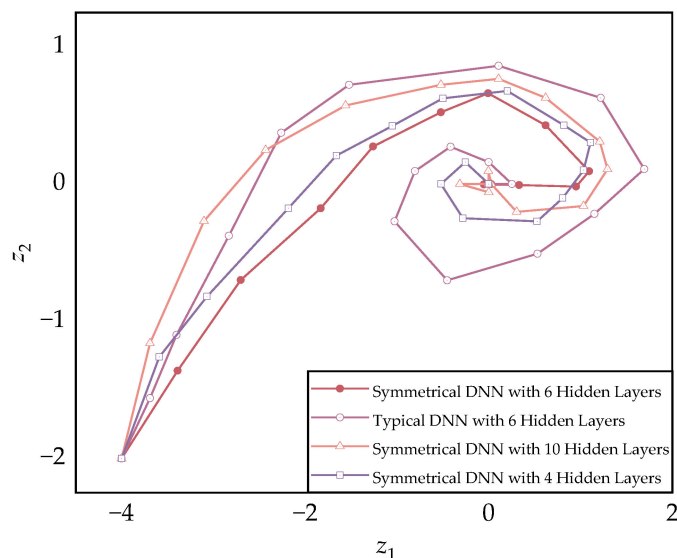
Figure 6. Nominal state trajectories under various DNN architectures.

The state trajectories for the proposed RMPC scheme are indicated in Figure 7. The solid line represents the state trajectory of the nominal system (3), while the dash-dot line is the state trajectory of the actual system (1). The error tube $\alpha_k E$ is depicted by green polytopes, while the 0-step homothetic tube controllability set X_0 is represented by the dark gray area. Obviously, the local state at each instance is regulated in an error tube $\alpha_k E$ centered around the trajectory of the nominal state. As anticipated, the cross-section of the error tube diminishes as the nominal state converges towards the origin.
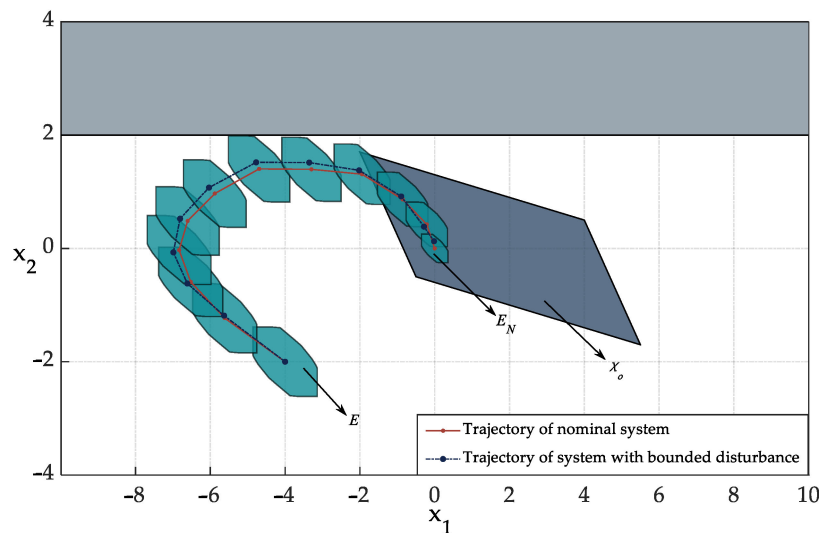
Figure 7. The state trajectories of the proposed algorithm ($N = 12$). Colors in the figure represent specific categories as follows: the green polytopes depict the error tube for every sampling times; the dark gray area represents the 0-step homothetic tube controllability set X_0; the gray area declares the undesirable state area.

Then, in order to make the comparison between the control performance of Algorithm 1 and the RMPC algorithm more apparent, N is set to 25. Figure 8 shows the state curves for Algorithm 1 and the HTMPC strategy. The state constraint is shown in the gray region. Algorithm 1 makes that initiating from an initial condition significantly distant from the desired equilibrium point enables faster convergence to the target state while maintaining a narrower range of fluctuation in state error when satisfying origin constraints for disturbances and state.
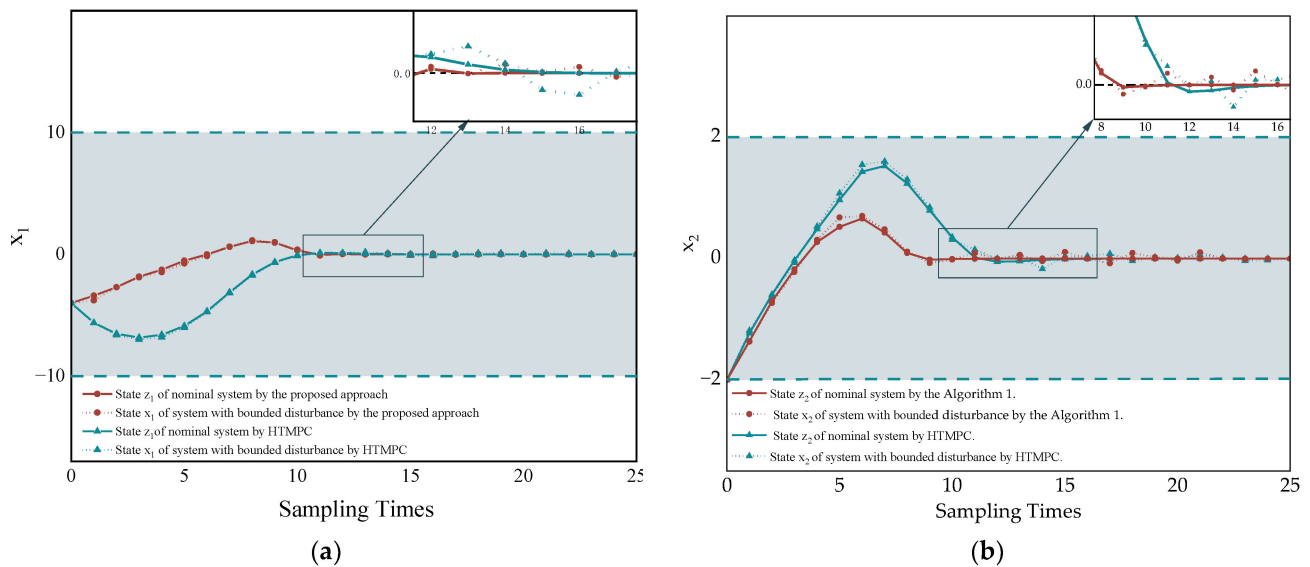
Figure 8. State astringency comparison between Algorithm 1 and the HTMPC ($N = 25$). (**a**) Curves of x_1 obtained by implementing two control algorithms, respectively; (**b**) curves of x_2 obtained by implementing two control algorithms, respectively.

Figure 9 presents the control input curves generated by two optimization methods. The region shown in gray is $\mathbb{U}$. Obviously, the control action of the actual system (1) consistently satisfies the control constraint. Meanwhile, Algorithm 1 accelerates the convergence of the control input toward the desired equilibrium point with reduced overshoot.
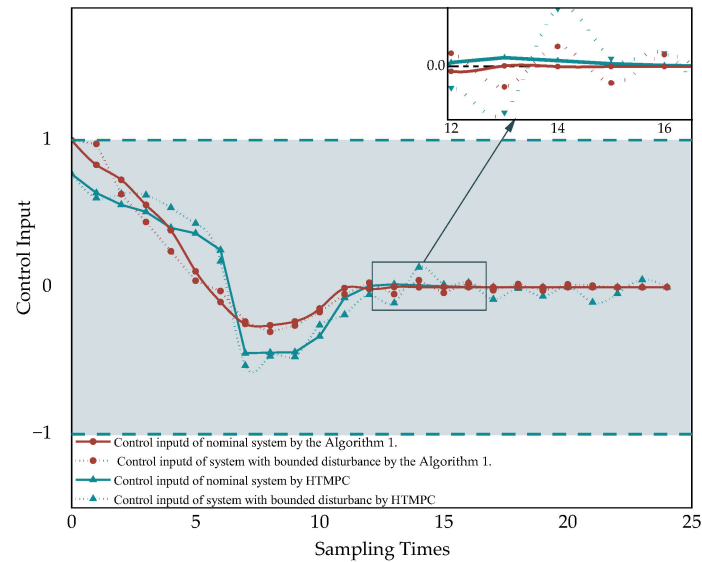
Figure 9. Control input astringency comparison between Algorithm 1 and the HTMPC ($N = 25$).

For the purpose of validating the efficacy of Algorithm 1 in reducing optimization time, a statistical analysis was conducted on the optimization time. Furthermore, to investigate the trend of optimization time, a slightly larger value of N ($N = 50$) was selected during the experimentation. As shown in Figure 10, the computational efficiency of Algorithm 1 is generally 2–3 orders of magnitude faster than HTMPC. In addition, as N increases, the calculation time for HTMPC exhibits an exponential growth trend. In contrast, the calculation time required by Algorithm 1 shows a gradual slowing trend and eventually stabilizes within 0.16 ms. Specifically, Algorithm 1 saves an average of 339.54 times more optimization time than HTMPC. When $N = 50$, A1 can save 726.23 times the optimization time compared to HTMPC.
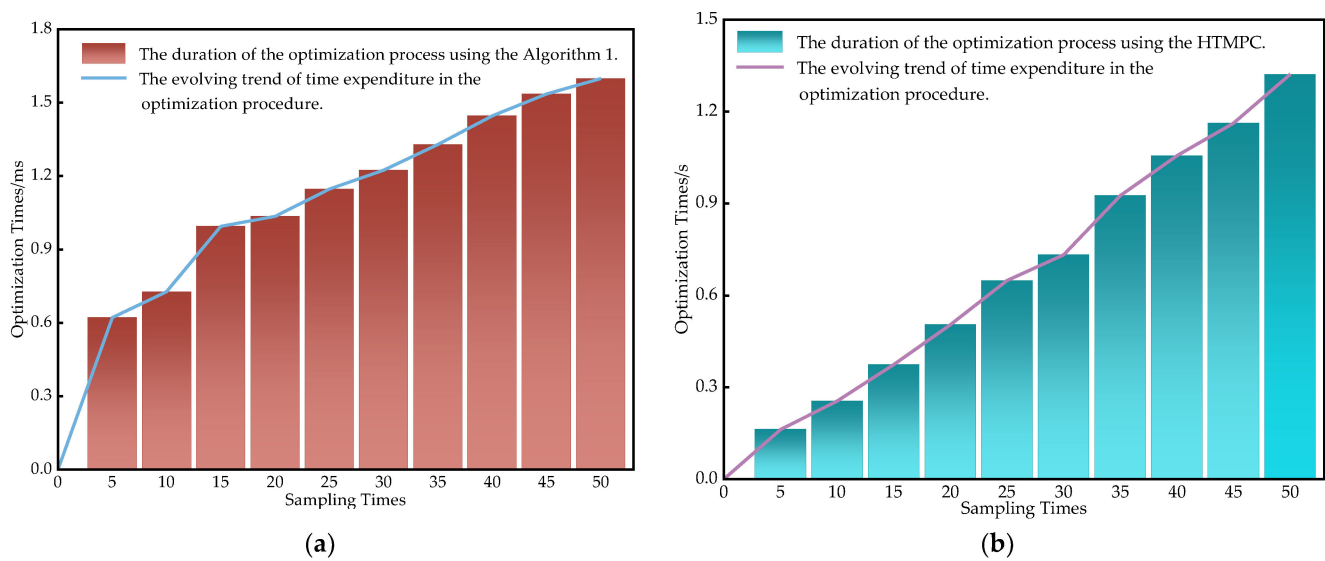
(a)

(b)

Figure 10. Comparison of Algorithm 1 and the HTMPC for computational efficiency ($N = 50$). (**a**) The statistical of computational time for Algorithm 1. (**b**) The statistical of computational time for HTMPC.

Example 2. *To further authenticate the proposed approach, consider the system of the form (1) with four state dimensions and two control input dimensions as*

$$A = \begin{bmatrix} 1 & 1.5 & 0 & 0 \\ 0.5 & -0.5 & 1 & 0 \\ 0 & 0.1 & 0.1 & 0 \\ 0.5 & 0 & 0.5 & 0.5 \end{bmatrix}, \ B = \begin{bmatrix} 0 & 1 \\ 1 & 0.1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{54}$$

Constraints are given by the inequalities as

$$x \in \mathbb{X} \triangleq \left\{ x \big| |x| \leq \begin{bmatrix} 5 \\ 5 \\ 2 \\ 2 \end{bmatrix} \right\}, \ u \in \mathbb{U} \triangleq \left\{ u \big| |u| \leq \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \tag{55}$$

The parameters are set to horizon $N = 30$, weighting matrices $Q_e = diag\{10, 10, 1, 1\}$ and $Q_v = diag\{0.01, 0.01\}$. The system is simulated according to the provided initial condition $x_0 = z_0 = (-3, -4, -1.5, 1)$. The other parameters of the system are under the same conditions as those in Example 1. Algorithm 1 will be implemented in this system to test its control performance for large-scale systems. Furthermore, the final DNN structure is determined by comparing the Euclidean norms of state errors generated when applying different deep neural network architectures, as illustrated in Figure 11. Specifically, the chosen DNN configuration comprises a symmetric deep neural network with eight hidden layers, each containing 14 neurons.

The Euclidean norm $\|e\|_2 = \sqrt{\sum_{i=1}^{4} (e_i)^2}$ is employed to depict the trend of state error changes. As indicated in Figure 11, it is observed that when applying a symmetric neural network with eight hidden layers, the system's state error is generally more minor and converges within the neighborhood of zero more quickly.
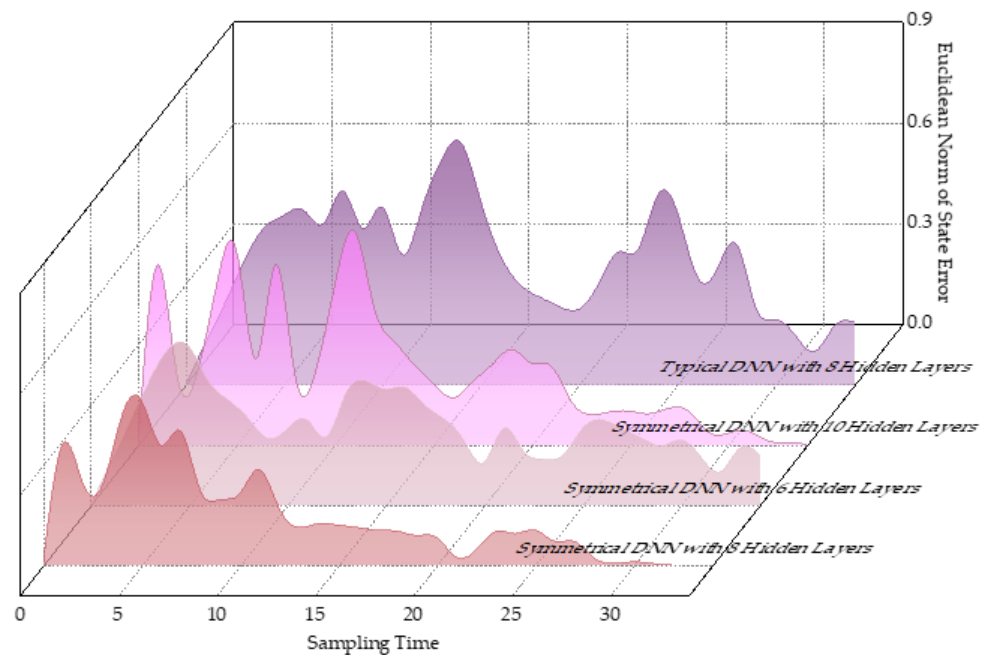
Figure 11. Euclidean norm of state error under various DNN architectures.

Figure 12 depicts the state variable curves for each dimension. The figure demonstrates that the time-varying nominal system obtained by online learning results in a slight error and shorter adjustment time during the convergence of the nominal state. The translucent area in these figures represents the range of error fluctuations; evidently, Algorithm 1 generally yields a bound on state errors than HTMPC, indicating greater flexibility in scaling the state tube. Furthermore, in Table 2, a visual comparison is performed using specific data to effectively demonstrate the error-constraining capabilities when evaluating the tracking performance of the actual system against the nominal system, employing Algorithm 1 and HTMPC. In order to mitigate the extreme influence of outliers, we opted for the mean squared error (MSE), known for its numerical stability, as the metric for assessing the tracking performance.

The utilization of Algorithm 1 for controlling a 4-D system, as illustrated in Table 4, leads to a minor MSE between the nominal and actual states across all four dimensions. The average MSE of the four dimensions is reduced by 67.86% when Algorithm 1 is employed, compared to its counterpart HTMPC. Consequently, the implementation of Algorithm 1 ensures a closer approximation of the actual state to the nominal state with reduced error.

Table 4. The MSE of state trajectories produced by two distinct control methodologies.

Control Strategy	Mean Squared Error			
	X_1	X_2	X_3	X_4
Algorithm 1	0.257407572	0.250064179	0.081418276	0.150326006
HTMPC	1.226454484	0.606667032	0.166021517	0.300930419

Moreover, the time-varying nominal system generated by Algorithm 1, as depicted in Figure 13, exhibits enhanced control input stabilization capabilities with a faster convergence rate and reduced overshoot.
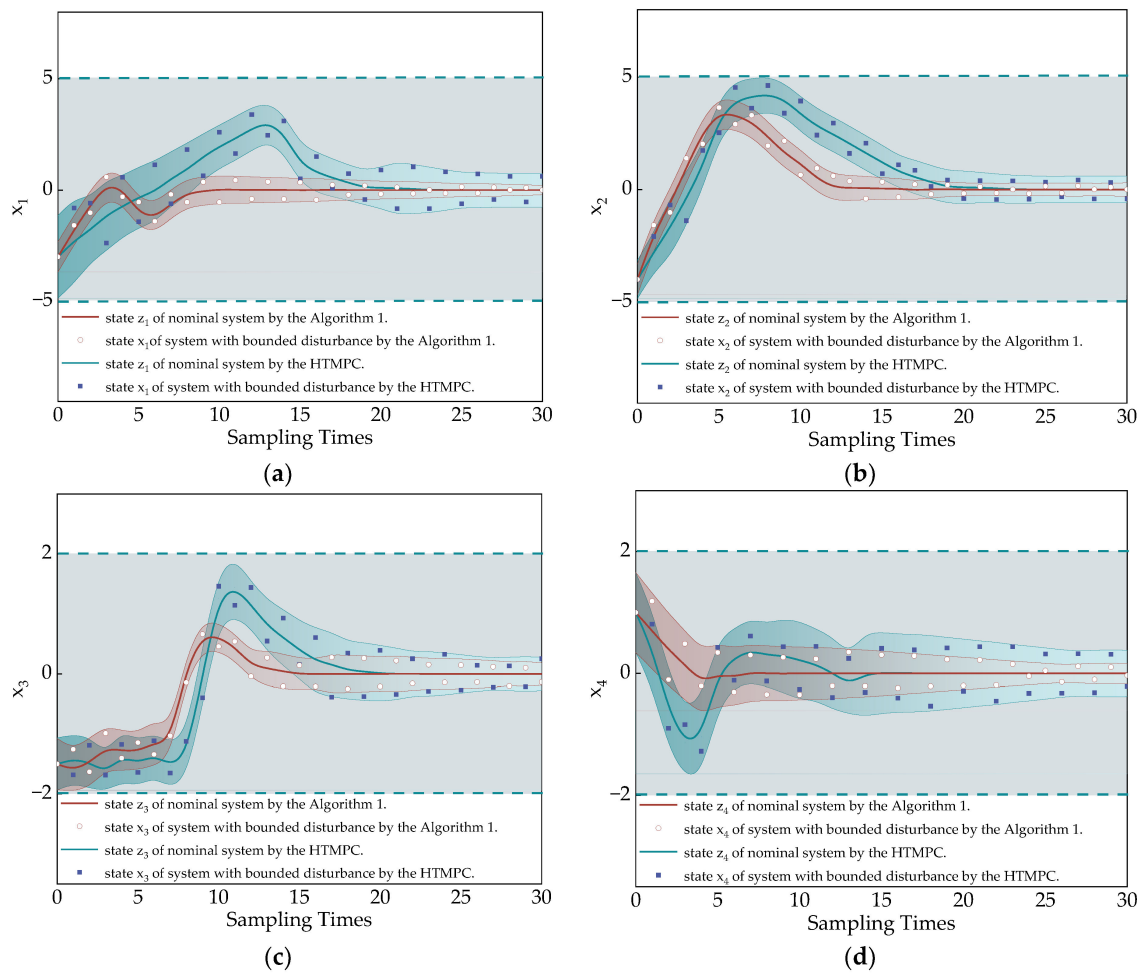
Figure 12. State astringency comparison between Algorithm 1 and the HTMPC ($N = 30$). (**a**) Curves of state obtained by implementing two distinct control algorithms, respectively. (**b**) Curves of state obtained by implementing two disti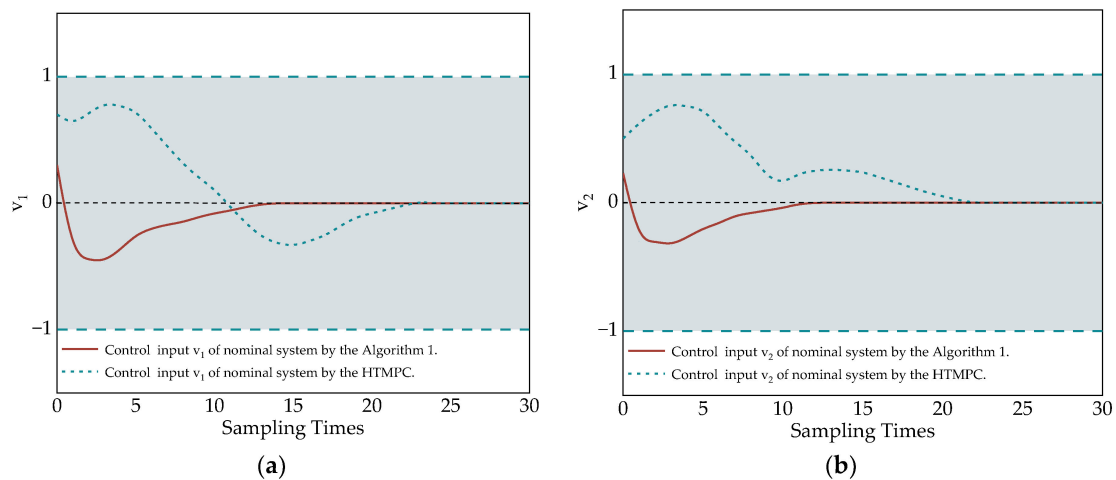nct control algorithms, respectively. (**c**) Curves of state obtained by implementing two distinct control algorithms, respectively. (**d**) Curves of state obtained by implementing two distinct control algorithms, respectively.

Figure 13. Control input astringency comparison between Algorithm 1 and the HTMPC ($N = 30$). (**a**) Curves of control input v_1 for the nominal system obtained by employing two distinct control algorithms, respectively. (**b**) Curves of control input v_2 for the nominal system obtained by employing two distinct control algorithms, respectively.

From a computational perspective, Algorithm 1 exhibits more pronounced advantages regarding computational efficiency for large-scale systems. As illustrated in Table 5, it can be observed that the proposed method significantly reduces the computation time to less than 6 ms when applied to four-dimensional input systems. In contrast, the HTMPC approach requires a longer computation time. On average, Algorithm 1 achieves optimization up to 7218.07 times faster than HTMPC.

Table 5. Comparison of Algorithm 1 and the HTMPC for computational efficiency.

Control Strategy	Horizon Length (N)				
	10	20	30	40	50
Algorithm 1	0.003938 s	0.004592 s	0.004823 s	0.004967 s	0.005094 s
HTMPC	23.179 s	27.674 s	30.239 s	38.098 s	49.837 s

The table denotes the calculation time unit "second" as "s".

5. Conclusions

This paper presents a mathematically rigorous and computationally tractable RMPC scheme for constrained linear systems with bounded disturbance. Firstly, a more flexible approach is proposed to adjust the size of the corresponding tube cross-section by incorporating a fuzzy-based tube size controller, which is influenced by both error magnitude and error variability ratio. Subsequently, the OCP for systems is reformulated as an online learning problem with iterative parameters. A time-varying nominal system for the control scheme is generated from the DNN-based nominal RMPC. Additionally, Dykstra's projection algorithm is incorporated into the DNN optimization process to ensure the feasibility of the successor state and control input. The proposed integrated control strategy significantly reduces the computational time while enhancing control effectiveness, thereby enabling its potential application in large-scale systems. Simulation results demonstrate the effectiveness of the proposed optimal control algorithm. The current study is constrained by the need for a measurable criterion for evaluating the suboptimal nature of the derived control law, thus impeding our ability to ascertain its degree of alignment with an optimal solution. To address this constraint, it might be imperative to devise metrics or algorithms capable of proficiently assessing the efficacy of the control.

References

1. Yang, H.Y.; Yan, Z.P.; Zhang, W.; Gong, Q.S.; Zhang, Y.; Zhao, L.Y. Trajectory tracking with external disturbance of bionic underwater robot based on CPG and robust model predictive control. *Ocean Eng.* **2022**, *263*, 112215. [CrossRef]
2. Shi, H.Y.; Li, P.; Su, C.L.; Wang, Y.; Yu, J.X.; Cao, J.T. Robust constrained model predictive fault-tolerant control for industrial processes with partial actuator failures and interval time-varying delays. *J. Process Control* **2019**, *75*, 187–203. [CrossRef]
3. Xie, Y.Y.; Liu, L.; Wu, Q.W.; Qian, Z. Robust model predictive control based voltage regulation method for a distribution system with renewable energy sources and energy storage systems. *Int. J. Electr. Power Energy Syst.* **2020**, *118*, 105749. [CrossRef]
4. Ojaghi, p.; Bigdeli, N.; Rahmani, M. An LMI approach to robust model predictive control of nonlinear systems with state-dependent uncertainties. *J. Process Control* **2016**, *47*, 1–10. [CrossRef]
5. Zheng, Y.P.; Li, D.W.; Xi, Y.G.; Zhang, J. Improved model prediction and RMPC design for LPV systems with bounded parameter changes. *Automatica* **2013**, *49*, 3695–3699. [CrossRef]

6.   Fleming, J.; Kouvaritakis, B.; Cannon, M. Robust tube MPC for linear systems with multiplicative uncertainty. *IEEE Trans. Autom. Control* **2015**, *60*, 1087–1092. [CrossRef]

7.   Hertneck, M.; Köhler, J.; Trimpe, S.; Allgöwer, F. Learning an approximate model predictive controller with guarantees. *IEEE Control Syst. Lett.* **2018**, *2*, 543–548. [CrossRef]

8.   Langson, W.; Chryssochoos, I.; Raković, S.V.; Mayne, D.Q. Robust model predictive control using tubes. *Automatica* **2004**, *40*, 125–133. [CrossRef]

9.   Mayne, D.Q.; Raković, S.V.; Findeisen, R.; Allgöwer, F. Robust output feedback model predictive control of constrained linear systems: Time varying case. *Automatica* **2009**, *45*, 2082–2087. [CrossRef]

10. Yu, S.Y.; Maier, C.; Chen, H.; Allgöwer, F. Tube MPC scheme based on robust control invariant set with application to Lipschitz nonlinear systems. *Syst. Control Lett.* **2013**, *62*, 194–200. [CrossRef]

11. Limon, D.; Alvarado, I.; Alamo, T.; Camacho, E.F. Robust tube-based MPC for tracking of constrained linear systems with additive disturbances. *J. Process Control* **2010**, *20*, 248–260. [CrossRef]

12. Cannon, M.; Kouvaritakis, B.; Raković, S.V.; Cheng, Q.F. Stochastic tubes in model predictive control with probabilistic constraints. *IEEE Trans. Autom. Control* **2011**, *56*, 194–200. [CrossRef]

13. Mayne, D.Q.; Seron, M.M.; Raković, S.V. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica* **2005**, *41*, 219–224. [CrossRef]

14. Mayne, D.Q.; Raković, S.V.; Findeisen, R.; Allgöwer, F. Robust output feedback model predictive control of constrained linear systems. *Automatica* **2006**, *42*, 1217–1222. [CrossRef]

15. Raković, S.V.; Kouvaritakis, B.; Findeisen, R.; Cannon, M. Homothetic tube model predictive control. *Automatica* **2012**, *48*, 1631–1638. [CrossRef]

16. Raković, S.V.; Cheng, Q.F. Homothetic tube MPC for constrained linear difference inclusions. In Proceedings of the Chinese Control and Decision Conference (CCDC), Guiyang, China, 25–27 May 2013; pp. 754–761.

17. Georgiou, A.; Tahir, F.; Jaimoukha, I.M.; Evangelou, S.A. Computationally Efficient Robust Model Predictive Control for Uncertain System Using Causal State-Feedback Parameterization. *IEEE Trans. Autom. Control* **2023**, *68*, 3822–3829. [CrossRef]

18. Zhang, B.; Yan, S. Asynchronous constrained resilient robust model predictive control for markovian jump systems. *IEEE Trans. Ind. Inform.* **2020**, *16*, 7025–7034. [CrossRef]

19. Olaru, S.; Dumur, D. A parameterized polyhedra approach for the explicit robust model predictive control. In *Informatics in Control, Automation and Robotics II*; Springer: Dordrecht, The Netherlands, 2007; pp. 217–226.

20. Lin, C.Y.; Yeh, H.Y. Repetitive model predictive control based on a recurrent neural network. In Proceedings of the 2012 International Symposium on Computer, Consumer and Control, Taichung, Taiwan, 4–6 June 2012; pp. 540–543.

21. Han, H.; Kim, H.; Kim, Y. An efficient hyperparameter control method for a network intrusion detection system based on proximal policy optimization. *Symmetry* **2022**, *14*, 161. [CrossRef]

22. Xu, K.; Huang, D.Z.; Darve, E. Learning constitutive relations using symmetric positive definite neural networks. *J. Comput. Phys.* **2021**, *428*, 110072. [CrossRef]

23. Di, M.M.; Forti, M.; Tesi, A. Existence and characterization of limit cycles in nearly symmetric neural networks. *IEEE Trans. Circuits Syst. I: Fundam. Theory Appl.* **2002**, *49*, 979–992.

24. Ma, C.Q.; Jiang, X.Y.; Li, P.; Liu, J. Offline computation of the explicit robust model predictive control law based on eeep neural networks. *Symmetry* **2023**, *15*, 676. [CrossRef]

25. Bumroongsri, P.; Kheawhom, S. An off-line robust MPC algorithm for uncertain polytopic discrete-time systems using polyhedral invariant sets. *J. Process Control* **2012**, *22*, 975–983. [CrossRef]

26. Lorenzen, M.; Cannon, M.; Allgöwer, F. Robust MPC with recursive model update. *Automatica* **2019**, *103*, 461–471. [CrossRef]

27. Moreno-Mora, F.; Beckenbach, L.; Streif, S. Performance bounds of adaptive MPC with bounded parameter uncertainties. *Eur. J. Control* **2022**, *68*, 100688. [CrossRef]

28. Bradtke, S.J. Reinforcement learning applied to linear quadratic regulation. *Adv. Neural Inf. Process. Syst.* **1992**, *5*, 295–302.

29. Wang, H.R.; Zariphopoulou, T.; Zhou, X.Y. Reinforcement learning in continuous time and space: A stochastic control approach. *J. Mach. Learn. Res.* **2020**, *21*, 8145–8178.

30. Schulman, J.; Levine, S.; Abbeel, P.; Jordan, M.; Moritz, P. Trust region policy optimization. *Int. Conf. Mach. Learn.* **2015**, *37*, 1889–1897.

31. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal policy optimization algorithms. *arXiv* **2017**, arXiv:1707.06347.

32. Kolmanovsky, I.; Gilbert, E.C. Theory and computation of disturbance invariant sets for discrete-time linear systems. *Math. Probl. Eng.* **1998**, *4*, 317–367. [CrossRef]

33. Wan, Z.; Pluymers, B.; Kothare, M.V.; De Moor, B. Efficient robust constrained model predictive control with a time varying terminal constraint set. *Syst. Control Lett.* **2006**, *55*, 618–621. [CrossRef]

34. Raković, S.V.; Kerrigan, E.C.; Kouramas, K.I.; Mayne, D.Q. Invariant approximations of the minimal robust positively invariant set. *IEEE Trans. Autom. Control* **2005**, *50*, 406–410. [CrossRef]

35. Nguyen, A.T.; Taniguchi, T.; Eciolaza, L.; Campos, V.; Palhares, R.; Sugeno, M. Fuzzy control systems: Past, present and future. *IEEE Comput. Intell. Mag.* **2019**, *14*, 56–58. [CrossRef]

36. Zeng, X.J.; Madan, G.S. Approximation accuracy analysis of fuzzy systems as function approximators. *IEEE Trans. Fuzzy Syst.* **1996**, *4*, 44–63. [CrossRef]

37. Bertsekas, D. Dynamic programming and optimal control: Volume I. *Athena Sci.* **2012**, *4*, 111–120.
38. Wu, L.G.; Su, X.J.; Shi, P. Dykstra's Algorithm for a constrained least-squares matrix problem. *Numer. Linear Algebra Appl.* **1996**, *3*, 459–471.
39. Sutton, R.S.; Mcallester, D.; Singh, S.; Mansour, Y. Policy gradient methods for reinforcement learning with function approximation. *Adv. Neural Info. Proc. Syst. (NIPS)* **2000**, *12*, 1057–1063.