



# Learning-based nonlinear model predictive control with accurate uncertainty compensation

Jingjie Xie · Xiaowei Zhao · Hongyang Dong 

Received: 24 October 2020 / Accepted: 5 May 2021 / Published online: 26 May 2021  
© The Author(s) 2021

**Abstract** A learning-based nonlinear model predictive control (LBNMPC) method is proposed in this paper for general nonlinear systems under system uncertainties and subject to state and input constraints. The proposed LBNMPC strategy decouples the robustness and performance requirements by employing an additional learned model and introducing it into the MPC framework along with the nominal model. The nominal model helps to ensure the closed-loop system's safety and stability, and the learned model aims to improve the tracking behaviors. As a core of the learned model construction, an online parameter estimator is designed to deal with system uncertainties. This estimation process effectively evaluates both the current and historical effects of uncertainties, leading to superior estimating performance compared with conventional methods. By constructing an invariant terminal constraint set, we prove that the LBNMPC is recursively feasible and robustly asymptotically stable. Numerical verifications for a two-link manipulator are conducted to validate the effectiveness and robustness of the proposed control scheme.

**Keywords** Nonlinear model predictive control · Learning-based control · Adaptive control · Parameter estimation

## 1 Introduction

Model predictive control (MPC) technique has received extensive attention over the recent decades [1–3], given its optimizing ability with respect to user-defined cost functions and its constraint handling ability regarding state and input constraints. Balancing the robustness and performance requirements is a challenging task in the design of advanced MPC schemes [4]. Though robust MPC methods [5], such as the min–max MPC [6] and tube-based MPC [7], can ensure robustness and constraint handling requirements, these methods usually over-prioritize robustness properties, causing performance degradation and conservativeness. Several control strategies for uncertain systems have been investigated recently [8–11]. Considering the efficiency and capability against uncertainties [12–14], adaptive MPC methods were proposed to reduce system conservativeness and improve the closed-loop performance in [15–17]. Moreover, a learning-based MPC (LBMPC) was proposed in [18] for linear systems. LBMPC combines the advantages of both the adaptive control and robust MPC, making it possible to improve performance under system uncertainties while guaranteeing robustness and safety requirements. The main principle of LBMPC is to minimize a cost function by using two parallel models: *i*) the learned model, which is modified online for performance enhancement purposes; *ii*) the nominal model, which is employed to

---

J. Xie · X. Zhao · H. Dong (✉)  
School of Engineering, University of Warwick,  
Coventry CV4 7AL, UK  
e-mail: Hongyang.Dong@warwick.ac.uk

ensure safety and stability [19]. This strategy, to some extent, decouples and balances the robustness and the performance of the closed-loop system. Given these merits, LBMPC has been employed in various practical applications [20–22].

However, existing LBMPC schemes [18–22] are only designed for linear nominal models. Moreover, only conventional certainty-equivalence (CE)-based adaptive control [23] or statistic-based estimation methods [24] were employed in the existing LBMPC frameworks. It is well understood that adaptive controllers synthesized through the CE principle cannot guarantee the convergence of parameter estimation errors to zero unless reference signals additionally satisfy the strong persistent-excitation (PE) conditions [25]. Therefore, the CE-based adaptive law may lead to poor transient performance and fail to estimate the true values of unknown parameters. Recently, the concurrent-learning adaptive control (CLAC) method was proposed to address the drawbacks of CE-based adaptive controllers [26]. State in a nutshell, this design innovatively uses specially selected and online recorded data concurrently with instantaneously incoming measurements for adaptation. Thus, the CLAC strategy can effectively estimate the unknown parameters based on both the current and historical effects, resulting in superior estimation performance under relaxed excitation conditions. However, how to embed the CLAC technique into the MPC framework (especially for discrete models) is still an open problem.

Motivated by these facts, a learning-based nonlinear MPC (LBNMPC) scheme with a concurrent-learning estimator is proposed in this paper for uncertain nonlinear systems, subject to input and state constraints. To meet the robustness and performance requirements of the closed-loop system simultaneously, an additional learned model enriched by the learning-based uncertainty estimator is introduced into the MPC framework. The main contributions are in order:

- (1) In contrast to existing LBMPC schemes [18, 22] that only consider linear nominal models, this work deals with nonlinear nominal models, and the stability and robustness of the closed-loop system can still be strictly guaranteed by utilizing robust MPC theory. The terminal

penalty and inequality constraints force the system states in a terminal region, and the recursive feasibility of the system guarantees the asymptotic stability of the closed-loop system.

- (2) A novel concurrent-learning estimator is designed in a discrete-time form and introduced into the LBNMPC. By employing both instantaneous and historical state data (the historical state data are recorded online), this estimator ensures superior estimation performance compared with conventional methods. It guarantees the exponential convergence of parameter estimation errors, subject to the satisfaction of a relaxed excitation condition.

The remainder of this paper is organized as follows. Section 2 establishes the main framework of our LBNMPC algorithm. Section 3 analyzes the recursive feasibility and stability of the LBNMPC. Section 4 presents the adaptive estimation procedure with stability analysis. In Sect. 5, numerical simulations are illustrated to show the advantages of LBNMPC. Finally, the paper ends with some conclusions in Sect. 6.

*Notations.*  $\mathbb{R}_{\geq a}$  signifies a set of numbers that their real parts are greater than  $a$ . The operation  $\|x\|_Q^2$  means  $\|x\|_Q^2 = x^T Q x$ , where  $Q$  is a positive-definite matrix. The operator  $\mathbb{X} \oplus \mathbb{Y} = \{x + y | x \in \mathbb{X}, y \in \mathbb{Y}\}$  denotes the Minkowski sum, where  $x$  and  $y$  are elements in the sets  $\mathbb{X}$  and  $\mathbb{Y}$ , respectively. The operator  $\mathbb{X} \ominus \mathbb{Y}$  means  $\mathbb{X} \ominus \mathbb{Y} = \{z \in \mathbb{R}^n : z \oplus \mathbb{Y} \subseteq \mathbb{X}\}$  [27]. We also denote  $\text{int}(\mathcal{F})$  as the interior of the set  $\mathcal{F}$ . Also,  $\lambda_{\min}(\cdot)$  and  $\lambda_{\max}(\cdot)$  denote the minimum and maximum eigenvalues of corresponding matrices.

## 1.1 Definitions

A function  $\varpi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  is type- $\mathcal{K}$  if it is continuous, strictly increasing, and  $\varpi(0) = 0$ . Moreover, the function  $\varpi$  belongs to class  $\mathcal{K}_{\infty}$  if  $\varpi(s) \rightarrow \infty$  for  $s \rightarrow \infty$  [27]. Furthermore, a function  $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  is a  $\mathcal{KL}$  function if  $\beta(\cdot, t)$  is type- $\mathcal{K}$  for  $t \geq 0$ ,  $\beta(s, \cdot)$  is nonincreasing for  $s \geq 0$ , and  $\beta(s, t) \rightarrow 0$  as  $t \rightarrow \infty$ . Finally, a set is a  $\mathcal{C}$ -set if it is compact and convex, and it is a  $\mathcal{PC}$ -set once it contains the origin.

### 1.2 LBNMPC strategy

As the main contribution of this work, a LBNMPC framework is developed in this section for general nonlinear systems under system uncertainties. The LBNMPC scheme aims to minimize the quadratic objective function consisting of a stage cost and a terminal cost, subject to nonlinear dynamics, state constraints, input constraints, and a terminal inequality constraint.

### 1.3 Problem statement

A LBMPC method was proposed in Ref. [18] to improve the system performance while ensuring robustness. It considers a linear nominal model and a learned model with uncertainties. The linear nominal model is used to guarantee the stability of the system. An identification tool is employed for the learned model to improve performance. The LBMPC was implemented in heating, ventilation, and air-conditioning systems in Ref. [20] and the real-time control of quadrotor helicopters in Refs. [19, 21], and [22]. However, all these elegant results are built on linear nominal models. Considering that many practical systems have strong nonlinear properties, this paper aims to address the optimal control problem for nonlinear systems subject to multiple constraints and system uncertainties. The main challenge comes from how to make a balance between adaptability and robustness while strictly guaranteeing the closed-loop stability. Motivated by these facts, a LBNMPC scheme with a high-performance learning estimator is proposed in this paper to solve optimal control problems for uncertain nonlinear systems under constraints. Firstly, the optimal control problem is formulated as follows.

Consider a discrete-time nonlinear system as follows

$$x^+ = f(x, u) + w(x, u) \tag{1}$$

where  $x \in \mathbb{R}^n$  is the state vector,  $u \in \mathbb{R}^m$  is the control input vector, and  $x^+ \in \mathbb{R}^n$  denotes the successor state of  $x$ . Besides,  $f(\cdot)$  denotes the nominal model of the system which is assumed to be twice continuously differentiable, and  $w(\cdot)$  is the uncertainty dynamics of the system.

Without loss of generality, we assume the system's equilibrium  $(x_e, u_e)$  is the origin. The system in (1) is subject to the state and control constraints:  $x \in \mathbb{X}$  and  $u \in \mathbb{U}$ . Here  $\mathbb{X} \subset \mathbb{R}^n$  and  $\mathbb{U} \subset \mathbb{R}^m$  are  $\mathcal{PC}$ -sets. Besides, the uncertainty dynamics is assumed to be bounded for  $x \in \mathbb{X}$  and  $u \in \mathbb{U}$ . Thus there exists a  $\mathcal{C}$ -set  $\mathbb{W}$  such that  $w(x, u) \in \mathbb{W}$  when  $x \in \mathbb{X}$  and  $u \in \mathbb{U}$ .

The LBNMPC is constructed to handle the stabilization issue of (1). As mentioned in introduction, a nominal model and a learned model are employed for controller design. The nominal model of (1) is described by

$$\bar{x}^+ = f(\bar{x}, \bar{u}) \tag{2}$$

where  $\bar{x}$  and  $\bar{u}$  are the state and input of the nominal model. The learned model is defined by

$$\hat{x}^+ = f(\hat{x}, \hat{u}) + \hat{w}(\hat{x}, \hat{u}) \tag{3}$$

where  $\hat{x}$  and  $\hat{u}$  are the induced state and control input based on the learned model, and  $\hat{w}$  denotes the estimation of  $w$ .

In our LBNMPC, an iterative optimization procedure is required to obtain the control sequence via optimizing a finite-horizon quadratic cost function, which is defined as follows with an initial state  $x_0$  and a desired state  $x_e$

$$V_N(x, k, u; x_0, x_e) \triangleq V_f(x(N); x_e) + \sum_{k=0}^{N-1} l(x(k) - x_e, u(k)) \tag{4}$$

where  $N$  is the prediction horizon,  $l(x, u)$  is the stage cost defined by  $l(x, u) \triangleq \|x\|_Q^2 + \|u\|_R^2$  with  $Q$  and  $R$  are positive-definite weight matrices. Besides, we denote

$$u(x, k; x_0, x_e) \triangleq \{u(x, 0; x_0, x_e), u(x, 1; x_0, x_e), \dots, u(x, N - 1; x_0, x_e)\}$$

as the control sequence for the whole prediction horizon. Moreover,  $V_f(\cdot)$  is the terminal cost function depicted as

$$V_f(x, x_e) = \|x - x_e\|_P^2 \tag{5}$$

where  $P$  is the terminal penalty matrix. We also define a terminal constraint set by

$$x(N) \in \Omega \tag{6}$$

It is designed to draw the states at the end of the finite prediction horizon to a neighborhood of the origin [28]. The construction process of this terminal constraint set  $\Omega$  is described in detail in the following subsection.

### 1.4 Construction of terminal constraint set

In Ref. [29], a novel robust MPC was proposed for a linear system with additive uncertainties to track changing targets. This controller has the ability to steer the uncertain system to a neighborhood of the target. Reference [30] designed MPC method for constrained systems with detailed stability and optimality analysis. Reference [31] introduced a robust MPC approach to guarantee the feasibility and robustness of linear systems under bounded disturbances and various constraints. Reference [32] presented a maximal output admissible set for linear MPC methods. Reference [33] proved the asymptotic closed-loop stability of the nonlinear MPC. These results provide fundamental design principles for the construction of terminal constraints and showcase how to ensure stability and feasibility of robust MPC controllers under disturbances and constraints.

Specifically, the terminal constraint set aims to block the move at the end of the prediction horizon and restrict the inherent behavior of the finite-horizon control. It is critical in providing stability, safety, robustness, and feasibility of MPC [34]. The terminal penalty matrix  $P$  and the terminal constraint set  $\Omega$  can be determined off-line. To this end, we consider the Jacobian linearization of the nominal dynamics at the equilibrium  $(x_e, u_e)$ :

$$\bar{x}^+ = A\bar{x} + B\bar{u} \tag{7}$$

where  $A = (\partial f / \partial \bar{x})(x_e, u_e)$  and  $B = (\partial f / \partial \bar{u})(x_e, u_e)$ .

Assuming this linearized system is controllable, then there exists a local linear feedback controller  $\bar{u} = K(\bar{x} - x_e) \in \mathbb{U}$  such that  $A^* = A + BK$  is asymptotically stable, where  $K$  is the feedback gain ( $K$  is a positive-definite matrix). Then, the terminal penalty matrix  $P$  and the matrix  $K$  can be determined by solving the following equation [28]

$$(A^*)^T P A^* - P = -(Q^* + \tau I) \tag{8}$$

where  $Q^* = Q + K^T R K$ , and  $\tau$  is a user-defined positive constant.

To guarantee the stability of the system, the invariant terminal constraint set  $\Omega$  is designed as

$$\Omega(x) = \{x | V_f(x, x_e) \leq \alpha\} \tag{9}$$

where  $\alpha$  is a constant and computed by the method proposed in Ref. [28].

*Remark 1* Note that the local linear feedback controller is only applied to calculate the terminal penalty matrix  $P$  and the terminal constraint set  $\Omega$  off-line and ensure the system asymptotic stability [33]. It is not directly employed to the actual control system. Besides, it should be emphasized that the terminal constraint set  $\Omega$  chosen by the linear feedback control is invariant for the nonlinear system under the MPC control law.

### 1.5 LBNMPC Strategy

The LBNMPC strategy proposed in this paper inspires from the tube MPC, a type of robust MPC, which can ensure the nonlinear system's real trajectory lies in a tube that surrounds the nominal trajectory. The width of the tube is restricted in a set  $\mathbb{C}$ , and the constraints set  $\mathbb{X}$  are shrunk by the width of the tube. Thereby, the nominal trajectory lies in  $\mathbb{X} \ominus \mathbb{C}$  and the real trajectory lies in  $\mathbb{X}$ . Similarly, for LBNMPC, the nominal and real trajectories lie in  $\mathbb{X} \ominus \mathbb{C}$  and  $\mathbb{X}$ , respectively.

Given the nominal model in (2) and learned model in (3), our LBNMPC is formulated as:

$$V_N(\hat{x}, k, \bar{u}; x_0, x_e) \triangleq V_f(\hat{x}(N); x_e) + \sum_{k=0}^{N-1} l(\hat{x}(k) - x_e, \bar{u}(k)) \tag{10}$$

$$\hat{x} = x, \bar{x} = x \tag{11}$$

$$\bar{x}^+ = f(\bar{x}, \bar{u}) \tag{12}$$

$$\hat{x}^+ = f(\hat{x}, \bar{u}) + \hat{w}(\hat{x}, \bar{u}) \tag{13}$$

$$\bar{x} \in \mathbb{X} \ominus \mathbb{C}, \bar{u} \in \mathbb{U} \ominus K\mathbb{C} \tag{14}$$

$$\bar{x}(N) \in \Omega \tag{15}$$

where  $\bar{u}$  is the control sequence and the first element of it (i.e.,  $\bar{u}$ ) is applied to both the nominal model and learned model. The set  $\mathbb{C}$  is defined as  $\mathbb{C}_0 = \{0\}$ ,  $\mathbb{C}_k = \oplus_{j=0}^{k-1} (A + BK)^j \mathbb{W}$  [18]. Also,  $\bar{x}(N) \in \Omega$  denotes the terminal constraints and  $\Omega$  is an invariant set restricted

by the local linear feedback controller [35]. In addition, Eq. (11) indicates that, after the optimal control input is applied to the system, the observed state  $x$  should be fed to both the nominal state  $\bar{x}$  and the learned state  $\hat{x}$  to construct the subsequent optimization problem.

*Remark 2* The terminal invariant set denotes a set that for all  $x \in \Omega$ , there exists a control  $u \in \mathbb{U}$  such that  $x^+ \in \Omega$  [36, 37]. Therefore, all trajectories of the system always stay in  $\Omega$  if they are starting from  $\Omega$ .

The solution of the above problem is denoted as  $\phi(k; x, \tilde{u})$  with the initial state  $x_0$  and the control sequence  $\tilde{u}$ . The state constraint  $\bar{x}(k) \in \mathbb{X} \ominus \mathbb{C}$ , control constraint  $\tilde{u}(k) \in \mathbb{U} \ominus \mathbb{K} \mathbb{C}$ , and terminal inequality constraint  $\phi(N, \bar{x}, \tilde{u}) \in \Omega$  lead to the control set  $U_N(\bar{x}, x_e)$  as follows

$$U_N(\bar{x}, x_e) \triangleq \{ \tilde{u} | \tilde{u}(k) \in \mathbb{U} \ominus \mathbb{K} \mathbb{C}, \bar{x}(k) \in \mathbb{X} \ominus \mathbb{C}, \forall k \in \{0, 1, \dots, N - 1\}, \phi(N, \bar{x}, \tilde{u}) \in \Omega \} \tag{16}$$

The optimal control problem  $P_N(x, k; x_0, x_e)$  is defined by

$$P_N(x, k; x_0, x_e) : V_N^0(x) = \min \{ V_N(\hat{x}, k, \tilde{u}; x_0, x_e) | \tilde{u} \in U_N(\bar{x}, x_e) \} \tag{17}$$

The initial solution of  $P_N(x, k; x_0, x_e)$  is

$$\tilde{u}^0(x, k, x_0, x_e) \triangleq \{ \tilde{u}^0(0; x, k, x_0, x_e), \tilde{u}^0(1; x, k, x_0, x_e), \dots, \tilde{u}^0(N - 1; x, k, x_0, x_e) \} \tag{18}$$

And the associated state sequence is

$$x^0(x, k, x_0, x_e) \triangleq \{ x^0(0; x, k, x_0, x_e), x^0(1; x, k, x_0, x_e), \dots, x^0(N - 1; x, k, x_0, x_e) \} \tag{19}$$

Then, the first element  $\tilde{u}^0(0; x, k, x_0, x_e)$  is applied to the LBNMPC. At the next sampling instant, this procedure operates repeatedly for the successor state.

*Remark 3* The difference between the LBNMPC and other MPC methods is that the LBNMPC is formulated based on two models (nonlinear nominal and learned models) simultaneously, which makes it

possible for dealing with uncertainties while preserving the properties of robust MPC.

*Remark 4* Note that in the LBNMPC, the cost function is constructed by the states of the learned model, while the constraints are imposed on the nominal model. This design ensures robustness when the learned model does not match the true dynamics.

*Remark 5* An important characteristic of LBNMPC is that the system safety, stability, and robustness are only related to the state, input, and terminal constraints based on the nominal model. Therefore, the safety & robustness requirement (guaranteed by the nominal model) and the performance enhancement (provided by the learned model) can be decoupled. As a result, the LBNMPC can make a trade-off between the system’s robustness and performance.

## 2 Stability analysis

In this section, the conditions for guaranteeing the stability of LBNMPC are presented in detail. It is noteworthy that the stability and robustness of the LBNMPC scheme are independent to the system uncertainties and the learning tools that are involved in the learned model. Some necessary definitions are given as follows for the subsequent stability analysis.

**Definition 1** [38], *Asymptotically Stable*: The system is said to be asymptotically stable (AS) about  $x_e$  on  $\mathcal{F} \subseteq \mathbb{R}^n$ , if there exists a type- $\mathcal{K}$  function  $\varpi$  such that for  $x_k \in \mathcal{F}$ , the condition  $|x_k - x_e| \leq \varpi(|x_k - x_e|, k)$  holds for  $k \geq 0$ .

**Definition 2** [39], *Robustly Asymptotically Stable*: The system is said to be robustly asymptotically stable (RAS) about  $x_e$  on  $\text{int}(\mathcal{F})$  with respect to measurement error (additive disturbance)  $e_k$ , if there exists a  $\mathcal{KL}$  function  $\beta$  and for each  $\varepsilon > 0$  and compact set  $\ell \subset \text{int}(\mathcal{F})$ , there exists  $\delta > 0$  such that for all the measurement errors  $e_k$  satisfying *i)*  $\max_k \|e_k\| < \delta$ ; *ii)*  $x_k \in \ell$  and  $\|x_k - x_e\| \leq \beta(\|x_k - x_e\|, k) + \varepsilon$  for all  $k \geq 0$ .

*Remark 6* The main difference between AS and RAS is that the RAS considers the measurement error (additive disturbance), and it guarantees the asymptotic stability of system with respect to the disturbance.

**Definition 3** [40], *Persistent Excitation*: A bounded signal  $g(\cdot): \mathbb{R} \rightarrow \mathbb{R}^{n \times m}$  is said to be persistent exciting (PE) if there exist positive constants  $a$  and  $\gamma$  such that for arbitrary  $t \geq 0$ , one has  $\int_t^{t+\gamma} g(\tau)g^T(\tau)d\tau \geq aI$ .

**Definition 4** [41], *Finite Excitation*: A bounded signal  $g(\cdot): \mathbb{R} \rightarrow \mathbb{R}^{n \times m}$  is said to be finite exciting (FE) if there exist positive constants  $a$  and  $\gamma$  such that in time interval  $[t, t + \gamma]$ , one has  $\int_t^{t+\gamma} g(\tau)g^T(\tau)d\tau \geq aI$ .

*Remark 7* The main difference between PE and FE is that PE requires the signal to be excited over the whole time, while FE just requires the signal to be excited over a finite-time interval. PE means the satisfaction of FE over the whole time. The PE condition is much stronger than the FE condition from the standpoint of practical engineering.

2.1 Stability of open-loop system

The trajectory generated by (12) with feasible solutions can satisfy the terminal inequality constraint in (15) and ensure the boundedness of the objective function in (10). Therefore, feasibility at each time step should be analyzed. We assume that the feasible point is  $p_n$ ; the system state and input predicted by the nominal model are denoted as  $\bar{x}(x, k, x_0, x_e, p_n) \in \mathbb{X} \ominus \mathbb{C}$ ,  $\bar{u}(x, k, x_0, x_e, p_n) \in \mathbb{U} \ominus K\mathbb{C}$ , respectively. By introducing the uncertain modeling error (denoted as  $d_n \in \mathbb{W}$ ) into the nominal model,  $\bar{x}^+ = f(\bar{x}, \bar{u}) + d_n$  should be considered for stability analysis. For the next feasible point  $p_{n+1}$ , we have  $\bar{x}(x, k, x_0, x_e, p_{n+1}) = \bar{x}(x, k, x_0, x_e, p_n) + d_{n+1}$  and  $\bar{u}(x, k, x_0, x_e, p_{n+1}) = \bar{u}(x, k, x_0, x_e, p_n) + Kd_{n+1}$ ,  $d_{n+1} \in \mathbb{W}$ . Thus,  $\bar{x}(x, k, x_0, x_e, p_{n+1}) \in \mathbb{X} \ominus (\mathbb{C} \oplus \mathbb{W}) \oplus \mathbb{W}$ , which implies  $\bar{x}(x, k, x_0, x_e, p_{n+1}) \in \mathbb{X} \ominus \mathbb{C}$ . Similarly, we have  $\bar{u}(x, k, x_0, x_e, p_{n+1}) \in \mathbb{U} \ominus K\mathbb{C}$ . As a result, for the next feasible point  $p_{n+1}$ , the associated state and input are still feasible.

**Lemma 1** *Considering a set of states  $\mathcal{X}_N$  satisfying (20) that leads to at least one control sequence  $\bar{u}$  meeting the state, control, and terminal constraints. Then, for the nominal system, the feasibility of the open-loop optimal control problem at  $k = 0$  implies its feasibility for all  $k > 0$ .*

$$x \in \mathcal{X}_N \triangleq \{x | U_N \neq \emptyset\} \tag{20}$$

*Proof* Under the condition that there exists an optimal control sequence  $\bar{u}^*(x, k, x_0, x_e)$  for the optimal control problem  $P_N(x, k; x_0, x_e)$  with the associated state sequence  $\bar{x}^*(x, k, x_0, x_e) \in \mathcal{X}_N$  at  $[k, k + N]$ , the state  $\bar{x}^*(x, k + N, x_0, x_e)$  belongs to the terminal constraint set  $\Omega$  due to the system feasibility. For the next optimal control problem  $P_N(x, k + \sigma; x_0, x_e)$  ( $\sigma > 0$  is a small sampling step), the initial state satisfies  $\bar{x}^0(x, k + \sigma, x_0, x_e) = \bar{x}^*(x, k + \sigma, x_0, x_e)$ . Then, a candidate control sequence  $\bar{u}(x, k + \sigma, x_0, x_e)$  under the local linear feedback controller for the problem  $P_N(x, k + \sigma; x_0, x_e)$  at  $[k + \sigma, k + \sigma + N]$  can be chosen as.

$$\begin{aligned} &\bar{u}(x, k + \sigma, x_0, x_e) \\ &= \begin{cases} \bar{u}^*(x, k, x_0, x_e) \text{ for } [k + \sigma, k + N] \\ K\bar{x}^0(x, k + \sigma, x_0, x_e) \text{ for } (k + N, k + \sigma + N] \end{cases} \end{aligned} \tag{21}$$

Note that the terminal constraints set  $\Omega$  is an invariant set restricted by the local linear feedback controller. Therefore, the initial state  $\bar{x}^0(x, k + \sigma, x_0, x_e) = \bar{x}^*(x, k + N, x_0, x_e) \in \Omega$  indicates that  $\bar{x}^0(x, k + \sigma + N, x_0, x_e) \in \Omega$  [42–44]. This completes the proof.

To sum up, for each prediction horizon of the optimal control problem in (10)–(15) with feasible solutions, the terminal penalty  $V_f(\hat{x}(N); x_e)$  considered in (10) and the terminal constraints in (15) can force the states at the end of the prediction horizon lie within a terminal region.

2.2 Stability of closed-loop system

We choose the cost function  $V_N(\hat{x}, k; x_0, x_e)$  as the Lyapunov function to analyze the closed-loop stability. Note that the cost function is related to the learned states at each prediction horizon. After solving the optimal control problem for each prediction horizon, the state  $x^0(x, k + 1, x_0, x_e)$  for next prediction horizon can be acquired by the real system model based on the solution of  $\bar{u}^*(x, k, x_0, x_e)$ . Then, we consider the following assumptions.

*Assumption 1* [45]:

- a)  $f(\cdot)$  is Lipschitz continuous in  $\mathbb{X} \times \mathbb{U}$ ,  $l(\cdot)$  and  $V_f(\cdot)$  are continuous.
- b)  $\Omega \subset \mathbb{X}$ ,  $\Omega$  is closed and

compact. c)  $u = k_f(x) \in \mathbb{U}, \forall x \in \Omega$ . d)  $f(x, k_f(x)) \in \Omega, \forall x \in \Omega$ .

**Assumption 2** [42]: For the stage cost  $l(\cdot)$  and the terminal cost  $V_f(\cdot)$ , there exist  $\mathcal{K}_\infty$  functions  $\alpha_l$  and  $\alpha_f$  satisfying

$$l(x, \tilde{u}) \geq \alpha_l(|x|), \forall x \in \mathcal{X}_N, \tilde{u} \in \mathbb{U} \tag{22}$$

$$V_f(x) \leq \alpha_f(|x|), \forall x \in \Omega \tag{23}$$

where  $x \in \mathcal{X}_N \triangleq \{x | U_N \neq \emptyset\}$  since the set  $U_N$  satisfies the state constraint, input constraint, and terminal constraint.

**Lemma 2** [45, 46]: *If the cost function satisfies the following condition: There exists a  $u = k_f(x) \in \mathbb{U}$  such that  $V_f(f(x, k_f(x)); x_e) + l(x - x_e, k_f(x)) \leq V_f(x, x_e)$  for all  $x \in \Omega(x_e)$ . Then, the closed-loop system is AS.*

Based on all these preliminaries, we propose the following theorem.

**Theorem 1** *The cost function  $V_N^0(\cdot)$  satisfies.*

$$V_N^0(x^+, k^+; x_0, x_e) \leq V_N^0(x, k; x_0, x_e) - l(x - x_e, k_N(x, k; x_0, x_e)) \tag{24}$$

and the closed-system system is AS.

*Proof* The solution of  $P_N(x, k; x_0, x_e)$  is.

$$\tilde{u}^0(x, k, x_0, x_e) \triangleq \{ \tilde{u}^0(0; x, k, x_0, x_e), \tilde{u}^0(1; x, k, x_0, x_e), \dots, \tilde{u}^0(N - 1; x, k, x_0, x_e) \} \tag{25}$$

The associated state sequence is

$$x^0(x, k, x_0, x_e) \triangleq \{ x^0(0; x, k, x_0, x_e), x^0(1; x, k, x_0, x_e), \dots, x^0(N; x, k, x_0, x_e) \} \tag{26}$$

and the first element  $\tilde{u}^0(0; x, k; x_0, x_e)$  is applied to the LBNMPC, and we denote  $x^+ = x^0(1; x, k, x_0, x_e)$ . Let  $\tilde{u}$  denote the following control sequence

$$\tilde{u}(x, k, x_0, x_e) \triangleq \{ \tilde{u}^0(1; x, k, x_0, x_e), \tilde{u}^0(2; x, k, x_0, x_e), \dots, k_f(x^0(N; x, k, x_0, x_e)) \} \tag{27}$$

which is feasible for  $P_N(x, k; x_0, x_e)$  but not necessarily optimal. Then, it follows that

$$\begin{aligned} V_N^0(x^+, k^+, \tilde{u}; x_0, x_e) &= V_N^0(x, k; x_0, x_e) \\ &\quad - l(x - x_e, k_N(x, k; x_0, x_e)) \\ &\quad - V_f(x^0(N; x, k; x_0, x_e)) \\ &\quad + l(f(x^0(N; x, k; x_0, x_e), \\ &\quad \quad k_f(x^0(N; x, k, x_0, x_e)))) \\ &\quad + V_f(f(x^0(N; x, k; x_0, x_e), \\ &\quad \quad k_f(x^0(N; x, k, x_0, x_e)))) \end{aligned} \tag{28}$$

Considering that  $V_N^0(x^+, k^+; x_0, x_e) = V_N^0(x^+, k^+, u; x_0, x_e) \leq V_N^0(x, k, \tilde{u}; x_0, x_e)$ , it can be obtained from (28) that

$$V_N^0(x^+, k^+; x_0, x_e) \leq V_N^0(x, k; x_0, x_e) - l(x - x_e, k_N(x, k; x_0, x_e)) \tag{29}$$

Therefore, the cost function is nonincreasing, and the closed-loop system is AS.

[18, 42, 45]. If there exists an additive disturbance  $w \in \mathbb{W}$ , then the asymptotic stability condition for the closed-loop system is modified as: There exists a  $u = k_f(x) \in \mathbb{U}$  such that

$$\begin{aligned} V_f(f(x, k_f(x)), w; x_e) + l(x - x_e, k_f(x)) \\ \leq V_f(x, w, x_e) + \delta, \forall w \in \mathbb{W} \end{aligned} \tag{30}$$

for all  $x \in \Omega(x_e)$ , where  $\delta \in (0, \infty)$ . Then, the LBNMPC is RAS with

$$\begin{aligned} V_N^0(x^+, k^+, w; x_0, x_e) \leq V_N^0(x, k, w; x_0, x_e) \\ - l(x - x_e, k_N(x, k; x_0, x_e)) + \delta, \forall w \in \mathbb{W} \end{aligned} \tag{31}$$

Therefore, the LBNMPC is RAS when *i)* Assumption 1 and Assumption 2 are satisfied; *ii)* the system uncertainty is bounded; *iii)* the terminal cost and the terminal invariant set force the state within the neighborhood of origin at the end of each prediction horizon; *iv)* the closed-loop system is feasible.

### 3 Concurrent-learning estimator

From Section III, the robustly asymptotic stability has been guaranteed with the nonlinear nominal model. However, the performance of LBNMPC requires the accurate estimation of the unmodeled dynamics. Some

examples to this end are given in [20, 47, 48]. In this section, a novel concurrent-learning-based estimation strategy is proposed and its convergence property is proved.

### 3.1 Estimator development

We aim to develop an estimator to compensate the uncertain dynamics:  $w(x, u)$ . In many applications,  $w(x, u)$  satisfies an affine representation, i.e., there exists a regressor matrix  $z(x, u) \in \mathbb{R}^{n \times p}$  and an unknown constant vector  $\phi \in \mathbb{R}^p$ , such that  $w(x, u) = z(x, u)\phi$ . And here  $p$  is the total number of unknown parameters. In fact, when  $w(x, u)$  cannot satisfy the affine representation, a single-layer neural network can be employed to reconstruct  $w(x, u)$ . Specifically, based on the Weierstrass approximation theorem [49],  $w(x, u)$  can be reconstructed by a very similar form with the affine representation:

$$w(x, u) = z(x, u)\phi + \varepsilon \quad (32)$$

But now  $z(x, u) \in \mathbb{R}^{n \times p}$  is a set of basis functions,  $\phi \in \mathbb{R}^p$  is a governed vector which contains the weights of corresponding basis functions, and  $\varepsilon$  is the reconstruction error. It has been well-understood that, by properly choosing a large enough set of basis functions, the reconstruction error  $\varepsilon$  is bounded in  $\mathbb{X} \times \mathbb{U}$  and can be arbitrarily small and negligible. Based on these facts, we assume the uncertain dynamics in our paper satisfies  $w(x, u) = z(x, u)\phi$ , and no matter  $z(x, u)$  is a regressor matrix or a set of user-defined basis functions. And the objective of the estimator is to evaluate the true value of  $\phi$ .

Based on  $w(x, u) = z(x, u)\phi$ , Eq. (1) can be rewritten as follows

$$x^+ = f(x, u) + z(x, u)\phi \quad (33)$$

Then, we design the following filtered variables:

$$x_f^+ = -\mu_f x_f + x \quad (34)$$

$$z_f^+ = -\mu_f z_f + z \quad (35)$$

$$f_f^+ = -\mu_f f_f + f \quad (36)$$

where  $\mu_f$  is the estimator coefficient satisfying  $0 < \mu_f < 1$ . Thus  $x_f^+$ ,  $z_f^+$  and  $f_f^+$  are bounded variables when  $x_f$ ,  $z_f$ , and  $f_f$  are bounded. Substituting (34)–(36) into (1) yields

$$(x_f^+)^+ - f_f^+ - z_f^+ \phi = -\mu_f (x_f^+ - f_f - z_f \phi) \quad (37)$$

We denote  $y = x_f^+ - f_f - z_f \phi$ , and then (37) indicates

$$y^+ = -\mu_f y \quad (38)$$

Since  $y$  is an exponentially vanishing term which converges to zero quickly, we have  $x_f^+ - f_f - z_f \phi \approx 0$ . Then, the estimation of  $\phi$  (denoted by  $\hat{\phi}$ ) can be divided into two terms

$$\hat{\phi} = \hat{\phi}_1 + \hat{\phi}_2 \quad (39)$$

And the successors of  $\hat{\phi}_1$  and  $\hat{\phi}_2$  are designed as

$$\hat{\phi}_1^+ - \hat{\phi}_1 = -\frac{k_I}{\lambda_Z} (z_f^T z_f \hat{\phi}_1 - z_f^T \beta) \quad (40)$$

$$\begin{aligned} \hat{\phi}_2^+ - \hat{\phi}_2 = & -\left( \frac{k_I}{\lambda_Z} \sum_{i=1}^q z_f^T(t_i) z_f(t_i) \right) \hat{\phi}_2 \\ & + \frac{k_I}{\lambda_Z} \sum_{i=1}^q z_f^T(t_i) \beta(t_i) \end{aligned} \quad (41)$$

where  $k_I$  is a user-defined positive constant,  $\beta = x_f^+ - f_f$ ,  $\lambda_Z = 1 + \lambda_{\max}(\varphi)$ , and  $\varphi = 1/\lambda_Z (z_f^T z_f + \sum_{i=1}^q z_f^T(t_i) z_f(t_i))$  is employed for ease of notation. Besides,  $t_i$  denotes a set of past time indexes with  $0 \leq t_i < t$ ,  $i = 1, 2, \dots, q$ , and here  $q$  is a constant which denotes the total number of historical data points. Based on (40) and (41), the estimator  $\hat{\phi}$  follows:

$$\begin{aligned} \hat{\phi}^+ - \hat{\phi} = & -\frac{k_I}{\lambda_Z} z_f^T z_f \hat{\phi} - \frac{k_I}{\lambda_Z} \left( \sum_{i=1}^q z_f^T(t_i) z_f(t_i) \right) \hat{\phi} \\ & + \frac{k_I}{\lambda_Z} z_f^T \beta + \frac{k_I}{\lambda_Z} \sum_{i=1}^q z_f^T(t_i) \beta(t_i) \end{aligned} \quad (42)$$

Equation (42) shows that not only real-time data but past measurements are concurrently introduced into the estimator, and the motivation of this design will be explained in the convergence analysis.

### 3.2 Convergence analysis

**Theorem 2** Consider the nonlinear system with uncertain dynamics as in (1), design the learning-based estimator as in (42). Then the estimation error

$\tilde{\phi} = \hat{\phi} - \phi$  is bounded. Moreover, if  $\lambda_{\min}(\varphi) > 0$ ,  $\tilde{\phi}$  exponentially converges to zero.

*Proof* Based on the fact that  $\beta = x_f^+ - f_f = z_f \phi$ , we have.

$$\begin{aligned} \hat{\phi}_1^+ - \hat{\phi}_1 &= -k_I(z_f^T z_f \hat{\phi}_1 - z_f^T \beta) = -k_I z_f^T (z_f \hat{\phi}_1 - z_f \phi) \\ &= -k_I z_f^T z_f \tilde{\phi} \end{aligned} \tag{43}$$

$$\begin{aligned} \hat{\phi}_2^+ - \hat{\phi}_2 &= -\frac{k_I}{\lambda_Z} \left( \sum_{i=1}^q z_f^T(t_i) z_f(t_i) \right) \hat{\phi}_2 \\ &\quad + \frac{k_I}{\lambda_Z} \sum_{i=1}^q z_f^T(t_i) \beta(t_i) \\ &= -\frac{k_I}{\lambda_Z} \left( \sum_{i=1}^q z_f^T(t_i) z_f(t_i) \right) \tilde{\phi}_2 \end{aligned} \tag{44}$$

Therefore, one has

$$\hat{\phi}^+ - \hat{\phi} = -k_I \varphi \tilde{\phi} \tag{45}$$

Then, consider the following storage function:

$$V = (1/k_I) \tilde{\phi}^T \tilde{\phi} \tag{46}$$

The successor value of  $V$  is  $V^+ = (\tilde{\phi}^+)^T \tilde{\phi}^+ / k_I$ . Accordingly,

$$\begin{aligned} V^+ - V &= (\hat{\phi}^+ - \phi)^T (\hat{\phi}^+ - \phi) - (\hat{\phi} - \phi)^T (\hat{\phi} - \phi) \\ &= (\hat{\phi}^+ - \phi)^T (\hat{\phi} - \phi) - (\hat{\phi} - \phi)^T (\hat{\phi} - \phi) \\ &\quad + (\hat{\phi}^+ - \phi)^T (\hat{\phi}^+ - \hat{\phi}) \\ &= (\hat{\phi}^+ - \hat{\phi})^T \tilde{\phi} + (\hat{\phi}^+ - \hat{\phi})^T (\hat{\phi}^+ - \hat{\phi}) + \tilde{\phi}^T (\hat{\phi}^+ - \hat{\phi}) \\ &= \tilde{\phi}^T (\varphi^2 - 2\varphi) \tilde{\phi} \end{aligned} \tag{47}$$

Since  $\varphi$  is positive-definite and  $\lambda_{\max}(\varphi) < 1$ , we have  $\varphi^2 - 2\varphi \leq 0$ . On this basis, Eq. (47) indicates  $V^+ - V \leq 0$ . So the estimation error  $\tilde{\phi}$  is bounded. Furthermore, if  $\lambda_{\min}(\varphi) > 0$ , then it is obvious that  $V^+ - V \leq -a \tilde{\phi}^T \tilde{\phi}$ , where  $a$  is a positive constant. Thus,  $\tilde{\phi}$  could exponentially converge to zero. The proof is complete.

*Remark 8* Unlike the conventional CLAC that relies on the accurate approximation of immeasurable variables (i.e., the variable  $x^+$  in this paper), the proposed estimator adopts filtered states and regressor matrices

in (34)–(36) and therefore circumvents the variable approximation requirements.

In summary, the detailed design procedure of the LBNMPC is described in this subsection. The detailed implementation procedure of the LBNMPC scheme is presented in Table 1, and the optimal control problem can be solved by the solver MOSEK [50]. For a better understanding of the proposed controller, the architecture of the LBNMPC is illustrated in Fig. 1 as well.

*Remark 9* In Theorem 2, we show that the estimation error  $\tilde{\phi}$  can exponentially converge to zero, subject to the satisfaction that  $\lambda_{\min}(\varphi) > 0$ . Recall the definition of FE conditions, this requirement can be guaranteed if  $z_f$  satisfies a FE condition. To ensure  $\varphi$  is full rank if  $z_f$  satisfies a FE condition, a simplest way is to add all incoming data of  $z_f$  into  $\varphi$  until  $\text{rank}(\varphi) = p$  (therefore,  $\lambda_{\min}(\varphi) > 0$ ). A more sophisticated method is to design a selection algorithm, and some examples are shown in [26, 33]. Moreover, without the past measurements  $\sum_{i=1}^q z_f^T(t_i) z_f(t_i)$  and  $\sum_{i=1}^q z_f^T(t_i) \beta(t_i)$ , one can only ensure  $\lim_{t \rightarrow \infty} z_f \tilde{\phi} = 0$  from (47). Under this condition,  $z_f$  is required to satisfy the PE condition as in Definition 3 to ensure the convergence of  $\tilde{\phi}$ . Note that this is a common requirement in conventional estimator/identifier designs. However, PE conditions are quite strong from the standpoint of practical engineering. In this paper, we relax the PE condition to the FE condition by employing past measurements.

#### 4 Applications to the control of a two-link manipulator

In this section, the effectiveness of the proposed LBNMPC scheme is validated via numerical simulations. A typical two-link robot manipulator model [51] is considered as follows

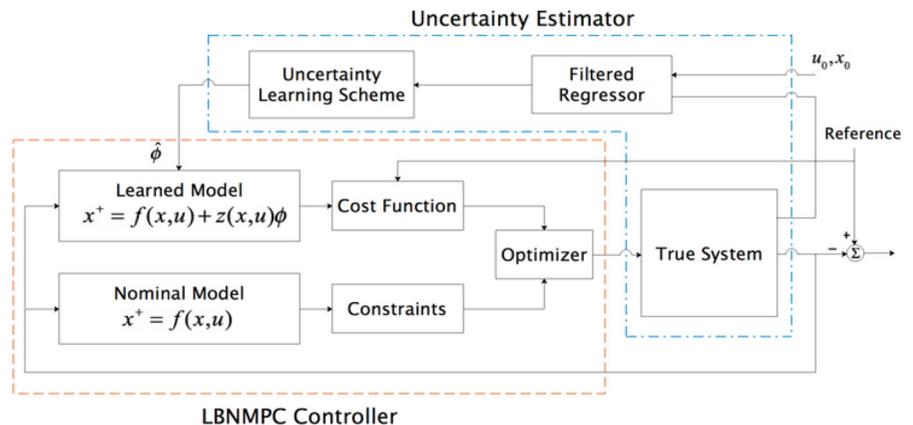
$$\begin{cases} \dot{x}_1 = x_2 \\ M(x_1, x_2) \dot{x}_2 + (V(x_1, x_2) + F(x_1, x_2)) x_2 + C(x_1, x_2) = u \end{cases} \tag{48}$$

where  $x_1 = [q_1 \ q_2]^T$  and  $x_2 = [q_3 \ q_4]^T$  denote the position and velocity vectors, respectively. Besides,  $u = [u_1 \ u_2]^T$  is the control torque vector, and

**Table 1** Architecture of LBNMPC

<b>Algorithm.</b> LBNMPC Implementation Procedure
<p><b>Offline:</b> (1) Compute the local linear feedback gain <math>K</math>.</p> <p>(2) Calculate the terminal penalty matrix <math>P</math> by solving (8).</p> <p>(3) Find an appropriate <math>\alpha</math> to determine terminal region.</p>
<p><b>Online:</b> For the iteration step (denoted as <math>iter</math>) and <math>k = 0, 1, \dots, N - 1</math>.</p> <p>(1) Set <math>iter = 1</math>. Select the initial values of <math>x, u, z, x_f, u_f, z_f</math> in the filtered dynamics, and the initial uncertainty estimation <math>\hat{\phi}</math>.</p> <p>(2) Construct the nominal model and the learned model.</p> <p>(3) Formulate the cost function and terminal constraint sets based on <math>P</math> and <math>\alpha</math>.</p> <p>(4) Employ the problem <math>P_N(x, k; x_0, x_e)</math> with state and control constraints.</p> <p>(5) Solve (10)-(15) by the optimal solver, obtain</p> $\bar{u}^0(x, k, x_0, x_e) \triangleq \{\bar{u}^0(0; x, k, x_0, x_e), \bar{u}^0(1; x, k, x_0, x_e), \dots, \bar{u}^0(N-1; x, k, x_0, x_e)\}.$ <p>(6) Apply the first control element <math>\bar{u}^0(0; x, k, x_0, x_e)</math> to the system and obtain the successor state <math>x^+</math>.</p> <p>(7) Update <math>\hat{\phi}^+</math> by the successor state via (42).</p> <p>(8) If <math>iter &lt; Max\_iter</math> and <math> x^+ - x_e  \geq tolerance</math>, set <math>iter = iter + 1</math> and return to step (2).</p>

**Fig. 1** The architecture of the LBNMPC with uncertainty learning



$$M = \begin{bmatrix} p_1 + 2p_3 \cos(q_2) & p_2 + p_3 \cos(q_2) \\ p_2 + p_3 \cos(q_2) & p_2 \end{bmatrix}$$

$$F = \begin{bmatrix} p_4 & 0 \\ 0 & p_5 \end{bmatrix} \quad C = \begin{bmatrix} p_6 \tanh(q_3) \\ p_7 \tanh(q_4) \end{bmatrix} V$$

$$= \begin{bmatrix} -p_3 \sin(q_2)q_4 & -p_3 \sin(q_2)(q_3 + q_4) \\ p_3 \sin(q_2)q_3 & 0 \end{bmatrix}$$

where the parameters  $p_1, p_2, p_3, p_4, p_5, p_6, p_7$  are set to 3.473, 0.196, 0.242, 5.3, 1.1, 8.45, and 2.35, respectively. The continuous-time system in (48) is discretized using Euler approximation as

$$\begin{cases} x_1(k+1) = T_s x_2(k) + x_1(k) \\ x_2(k+1) = T_s f(k) + T_s g(k)u(k) + x_2(k) \end{cases} \quad (49)$$

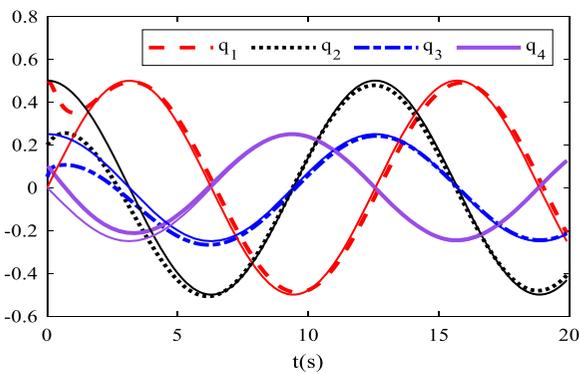
where  $T_s$  is the sampling period,  $f(k) = -M^{-1}[(V + F)x_2(k) + C]$ ,  $g(k) = -M^{-1}$ .

In this simulation case, we assume that the parameters  $p_5$  and  $p_7$  are unknown for controller design and thus they need to be estimated by the proposed learning procedure. The initial guesses of them are  $\hat{p}_5(0) = 0.6$  and  $\hat{p}_7(0) = 1.8$ . The parameter  $k_f$  in the

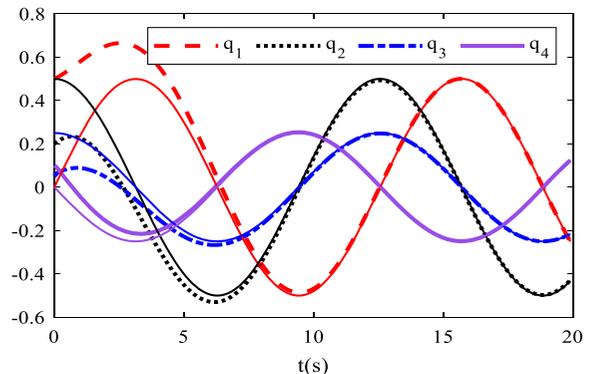
filtered regressor is selected as  $k_f = 0.2$ . The prediction horizon is set to  $N = 7$ , and the weighting matrices  $Q$  and  $R$  are chosen as  $Q = 3000 \times I_{4 \times 4}$  and  $R = I_{2 \times 2}$ . The constant  $\alpha$  is computed as  $\alpha = 1.34$ . Through the Jacobian linearization of the nominal system at the origin, the matrices  $A$  and  $B$  are calculated by (7) as

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -4.62 & 2.59 \\ 0 & 0 & 10.32 & -23.39 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0.34 & -0.75 \\ -0.75 & 6.78 \end{bmatrix}$$

The terminal penalty matrix  $P$  and the local linear feedback gain  $K$  are calculated by (8) as

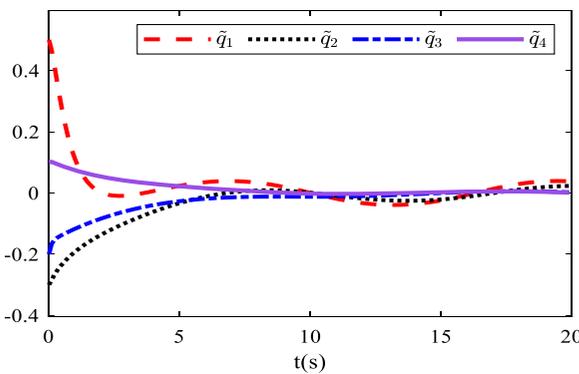


(a) Tracking trajectories by LBNMPC without uncertainty learning

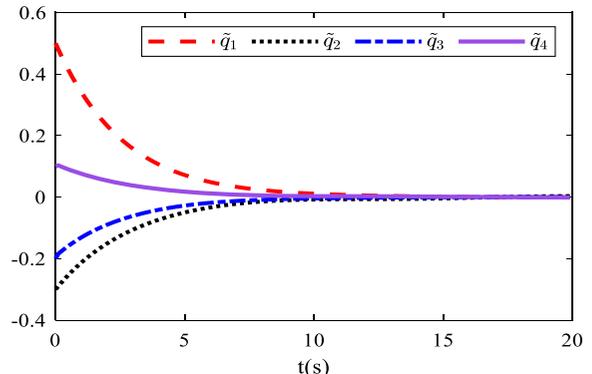


(b) Tracking trajectories by LBNMPC with uncertainty learning

Fig. 2 Tracking performance by LBNMPC with/without uncertainty learning



(a) Tracking errors by LBNMPC without uncertainty learning



(b) Tracking errors by LBNMPC with uncertainty learning

Fig. 3 Tracking errors by LBNMPC with/without uncertainty learning

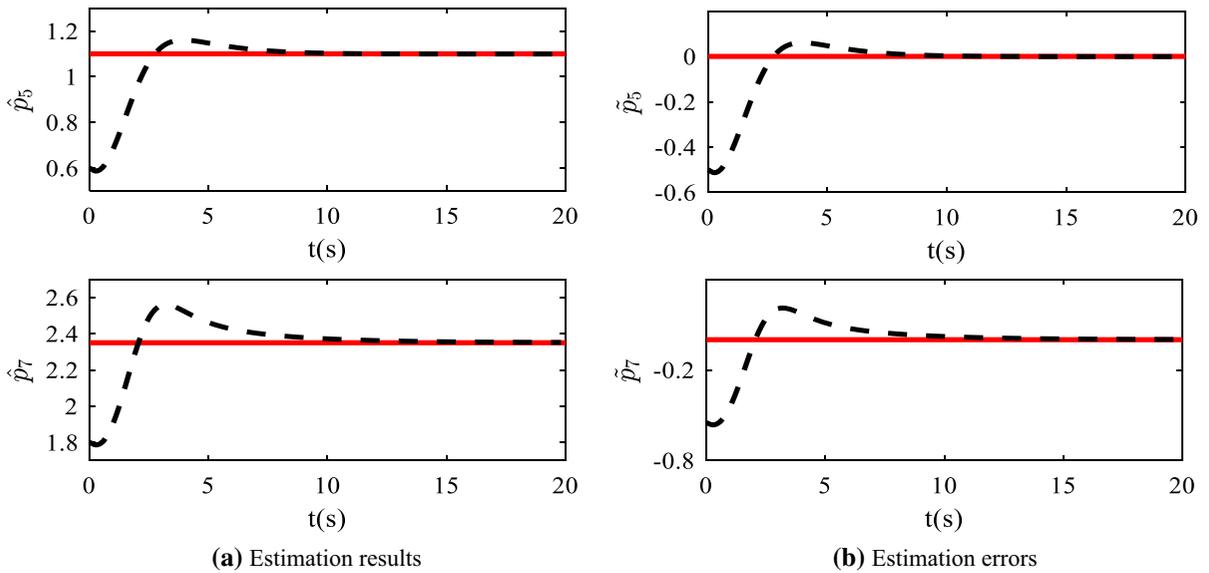


Fig. 4 Uncertainty estimation results

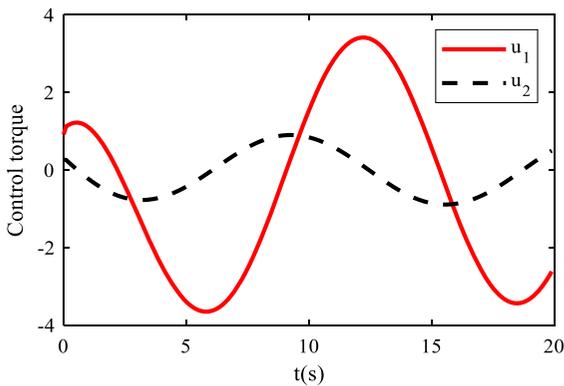


Fig. 5 Results for the control torque performance

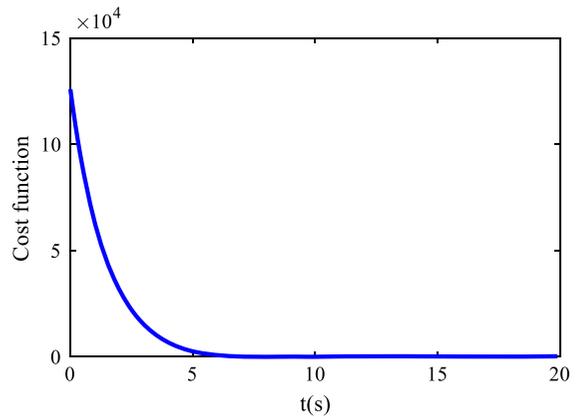


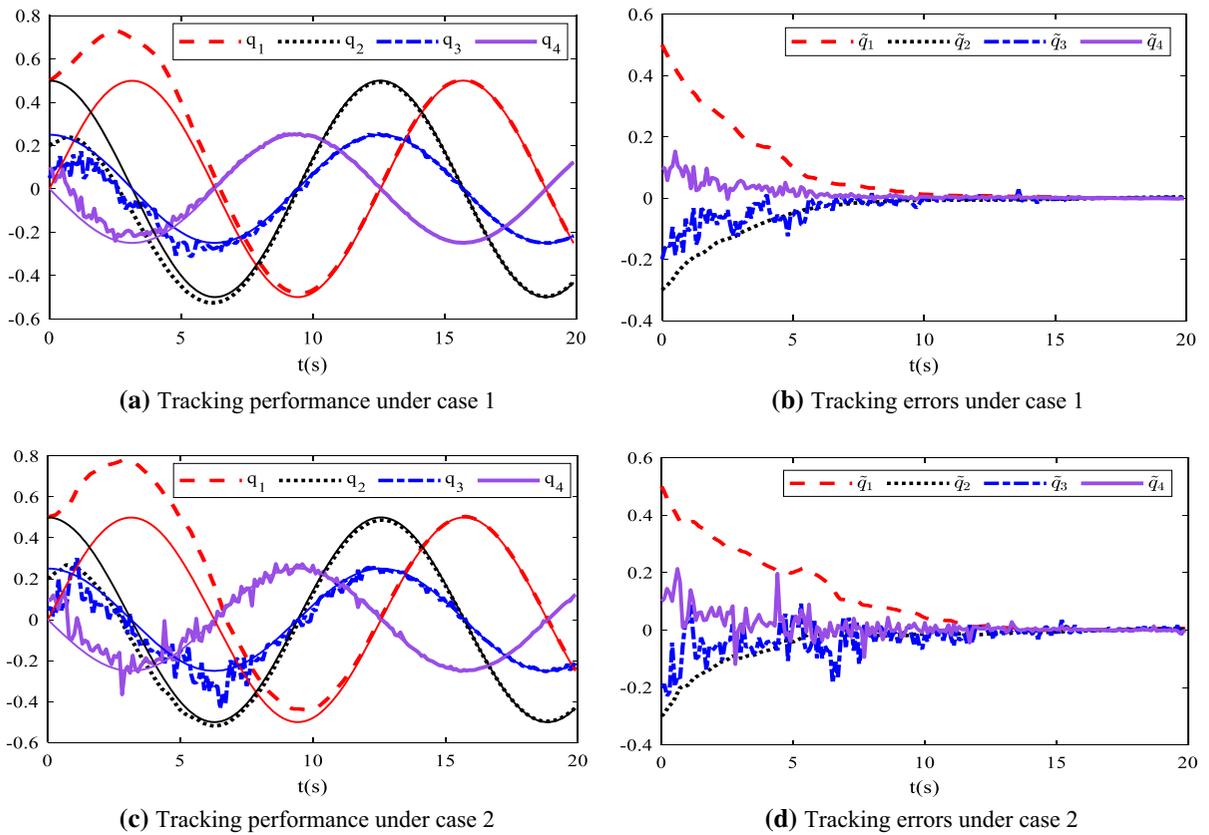
Fig. 6 The cost function variations

$$P = \begin{bmatrix} 307.70 & 0.73 & 6.86 & 0.76 \\ 0.73 & 300.39 & 0.72 & 0.34 \\ 6.86 & 0.72 & 6.49 & 0.72 \\ 0.76 & 0.34 & 0.72 & 0.33 \end{bmatrix} K \\
 = \begin{bmatrix} 173.20 & -1.02 & 163.89 & -0.60 \\ 1.02 & 173.20 & 1.46 & 169.98 \end{bmatrix}$$

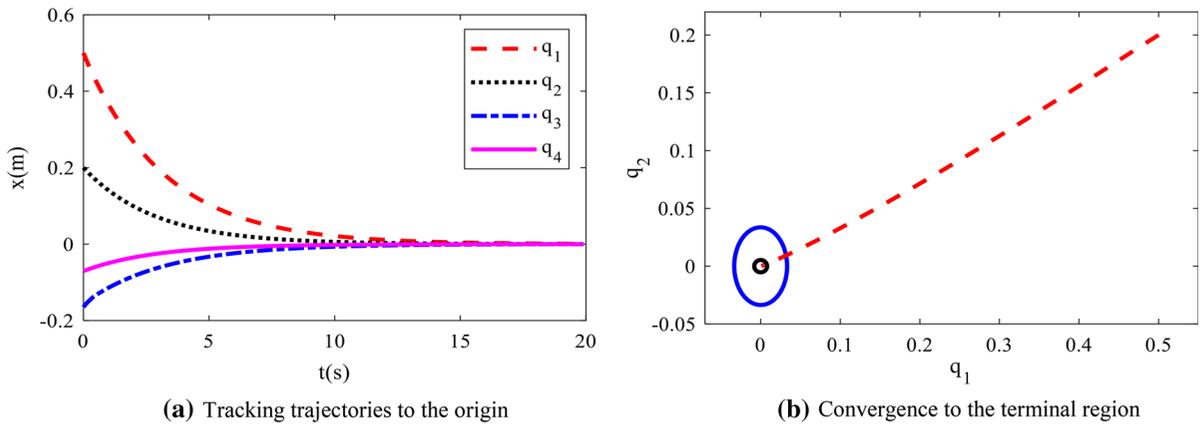
The set  $\mathbb{U}$  is selected as  $\mathbb{U} = [-5, 5]$ . The set  $\mathbb{W}$  is chosen via the algorithm in [52] that exploits  $\mathbb{W}$  based on the Taylor reminder theorem with a “safety-margin” [18]. Then, the set  $\mathbb{C}$  can be calculated by its definition ( $\mathbb{C}_0 = \{0\}$ ,  $\mathbb{C}_k = \bigoplus_{j=0}^{k-1} (A + BK)^j \mathbb{W}$  [18]). Therefore, the constraint in (14) can be determined, which ensures the closed-loop system’s safety. This

reflects the main characteristic of LBNMPC that decouples the system’s safety and the performance enhancement. The desired trajectory is  $x_d = [0.5 \sin(0.5k), 0.5 \cos(0.5k), 0.25 \cos(0.5k), -]$ , and the initial state is  $x_0 = [0.5, 0.2, 0.05, 0.1]^T$ .

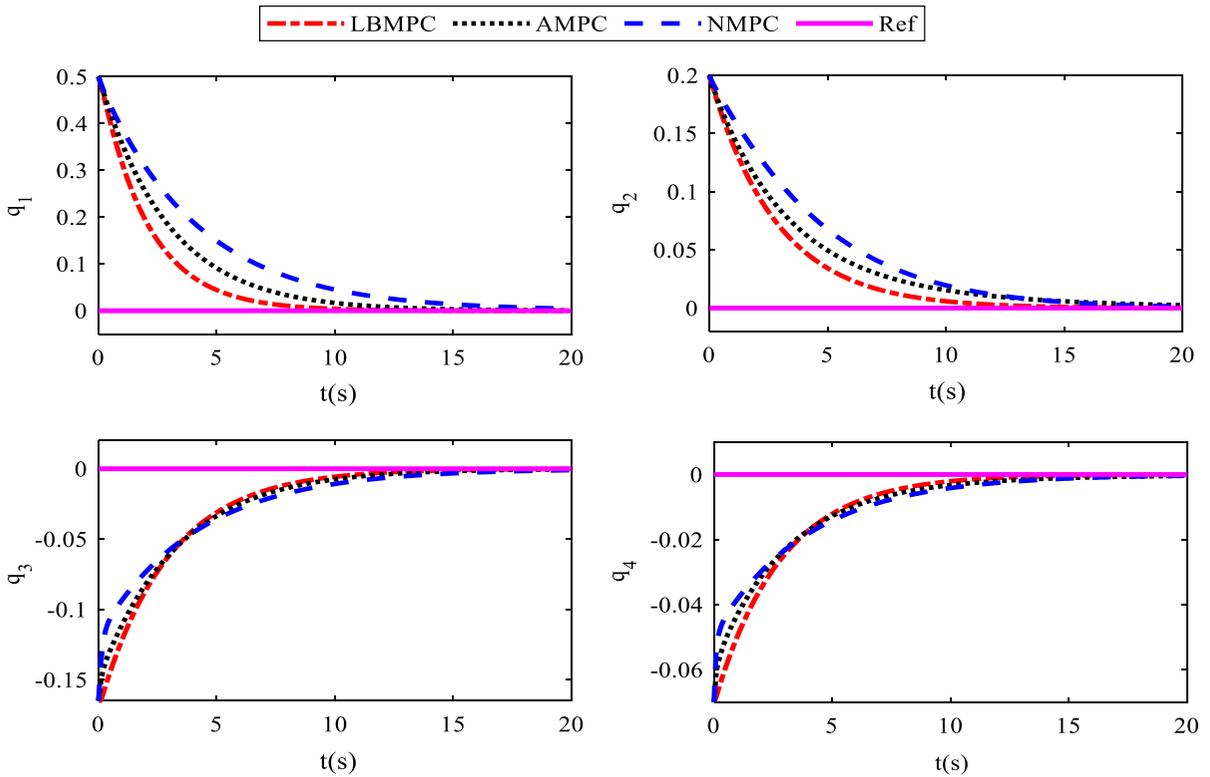
Under these conditions, the simulation results under LBNMPC without the estimation procedure are presented in Fig. 2a, while the results with the uncertainty learning method are shown in Fig. 2b. The thin solid lines in Fig. 2 represent the reference trajectories. It can be observed that the system states can successfully track the desired trajectories in both cases. Moreover, it can be seen from Fig. 3 that the LBNMPC leads to



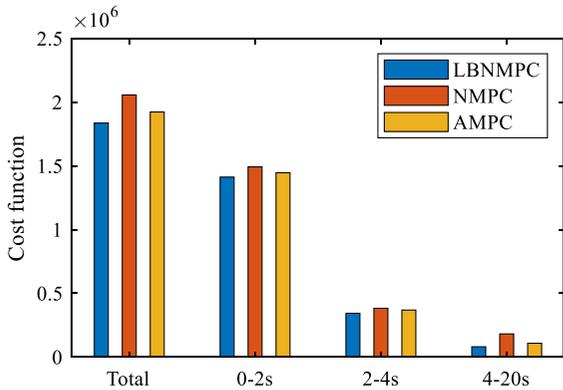
**Fig. 7** Tracking performance by LBNMPC under disturbances



**Fig. 8** Tracking performance under  $x_d = [0 \ 0 \ 0 \ 0]^T$



**Fig. 9** Comparative tracking trajectories under various MPC controllers



**Fig. 10** The values of cost function under different MPC controllers

superior performance than that of the one without parameter estimation.

The uncertainty learning performance for the unknown parameters  $p_5, p_7$  is presented in Fig. 4a. The solid lines in Fig. 4a depict the real values of  $p_5, p_7$ , and the dashed lines denote the estimated values

$\hat{p}_5, \hat{p}_7$ . It can be seen from Fig. 4a that the estimated parameters can converge to their real values. The estimation errors ( $\tilde{p}_5$  and  $\tilde{p}_7$ ) in Fig. 4b show that the proposed uncertainty learning scheme can ensure the estimation errors converge to zero.

The optimal control sequence is presented in Fig. 5. The control sequence in Fig. 5 satisfies the input constraints formulated in (14). The cost function variations are presented in Fig. 6. The cost function decreases gradually and finally converges to zero, reflecting the asymptotic stability of the LBNMPC approach.

In addition to the unknown parameters, the robustness of the LBNMPC is verified under external disturbances on velocity measurements. Two difference cases are considered here: *i*) Case 1: the velocity measurements are polluted by a zero-mean Gaussian noise with a standard deviation as  $1.0 \times 10^{-3}$ ; *ii*). Case 2: the velocity measurements are polluted by a zero-mean Gaussian noise with a standard deviation as  $3.0 \times 10^{-3}$ . The simulation results are shown in Fig. 7. Figures 7a and 7c show the tracking

trajectories under case 1 and case 2, respectively. The thin solid lines are the desired trajectories. Figures 7b and 7d show the corresponding tracking errors. It can be seen that the position states can precisely track the reference trajectories with good performance even under polluted velocity measurements. All these results show that the proposed LBNMPC approach can maintain good performance under disturbances.

To evaluate the properties of the proposed LBNMPC in detail, we also consider a stabilization case by setting  $x_d = [0 \ 0 \ 0 \ 0]^T$ . The results are given in Fig. 8. One can see that all the states can converge to the origin under the proposed controller. Figure 8b illustrates the tracking errors and the terminal region  $\Omega$ . It shows that the closed-loop system satisfies the constraints in (15), validating the stability and convergence properties of our LBNMPC.

Moreover, two different MPC approaches are also employed for comparison: *i*) the adaptive model predictive control (AMPC) with the proposed estimation procedure, and *ii*) the nonlinear model predictive control (NMPC) without uncertainty compensation. The comparative results of tracking trajectories and cost functions are presented in Figs. 9 and 10, respectively. It can be observed that both the LBNMPC and AMPC have good performance in tracking the desired trajectories, while the NMPC leads to larger tracking errors. Moreover, it can be seen from Fig. 10 that the proposed LBNMPC controller leads to a less cost than the AMPC and NMPC.

To sum up, the proposed LBNMPC is capable of achieving good performance, subject to uncertainties, disturbances, and various constraints.

## 5 Conclusion

A LBNMPC method was proposed in this paper to solve the optimal control problems of nonlinear systems subject to multiple constraints and system uncertainties. The control strategy was based on two models, i.e., the nonlinear nominal model and the learned model. The nominal model guarantees the stability and robustness of the LBNMPC, while the learned model improves the control performance via a novel concurrent-learning estimator. The key feature of our estimator is that it includes not only real-time data but also past measurements into the estimating

framework, achieving precise estimation under a relaxed excitation condition. We showed that our LBNMPC could decouple the robustness and performance and ensure the feasibility, stability, and convergence of the closed-loop system. Extensive simulations and comparative analyses illustrated that LBNMPC could lead to superior tracking performance and robustness compared with other methods.

**Funding** This work has received funding from the UK Engineering and Physical Sciences Research Council (Grant Number: EP/S001905/1).

## Declarations

**Conflicts of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Chai, R., Savvaris, A., Chai, S.: Integrated missile guidance and control using optimization-based predictive control. *Nonlinear Dyn.* **96**(2), 997–1015 (2019)
2. Zhao, J., Zhou, S., Zhou, R.: Distributed time-constrained guidance using nonlinear model predictive control. *Nonlinear Dyn.* **84**(3), 1399–1416 (2016)
3. Yao, P., Wang, H., Ji, H.: Gaussian mixture model and receding horizon control for multiple UAV search in complex environment. *Nonlinear Dyn.* **88**(2), 903–919 (2017)
4. Pin, G., Raimondo, D.M., Magni, L., et al.: Robust model predictive control of nonlinear systems with bounded and state-dependent uncertainties. *IEEE Trans. Autom. Control* **54**(7), 1681–1687 (2009)
5. Li, H., Shi, Y.: Robust distributed model predictive control of constrained continuous-time nonlinear systems: A Robustness Constraint Approach. *IEEE Trans. Autom. Control* **59**(6), 1673–1678 (2014)
6. Alessandri, A., Gaggero, M., Tonelli, F.: Min-max and predictive control for the management of distribution in supply chains. *IEEE Trans. Control Syst. Technol.* **19**(5), 1075–1089 (2011)

7. Abbas, H.S., Mannel, G., Hoffmann, C.H., et al.: Tube-based model predictive control for linear parameter-varying systems with bounded rate of parameter variation. *Automatica* **107**, 21–28 (2019)
8. Wang, C., Agarwal, R.P., O'Regan, D.: Calculus of fuzzy vector-valued functions and almost periodic fuzzy vector-valued functions on time scales. *Fuzzy Sets Syst.* **375**, 1–52 (2019)
9. Sakthivel, R., Wang, C., Santra, S., Kaviarasan, B.: Non-fragile reliable sampled-data controller for nonlinear switched time-varying systems. *Nonlinear Anal. Hybrid Syst* **27**, 62–76 (2018)
10. Wang, C., Agarwal, R.P.: Almost periodic solution for a new type of neutral impulsive stochastic Lasota-Ważewska timescale model. *Appl. Math. Lett.* **70**, 58–65 (2017)
11. Sakthivel, R., Joby, M., Wang, C., Kaviarasan, B.: Finite-time fault-tolerant control of neutral systems against actuator saturation and nonlinear actuator faults. *Appl. Math. Comput.* **332**, 425–436 (2018)
12. Hu, Q.: Robust adaptive sliding mode attitude maneuvering and vibration damping of three-axis-stabilized flexible spacecraft with actuator saturation limits. *Nonlinear Dyn.* **55**, (4), 301 (2009)
13. Zhang, C., Ma, G., Sun, Y., et al.: Prescribed performance adaptive attitude tracking control for flexible spacecraft with active vibration suppression. *Nonlinear Dyn.* **96**(3), 1909–1926 (2019)
14. Hu, Q., Shao, X., Guo, L.: Adaptive fault-tolerant attitude tracking control of spacecraft with prescribed performance. *IEEE/ASME Trans. Mechatron.* **23**(1), 331–341 (2018)
15. Zhang, K., Shi, Y.: Adaptive model predictive control for a class of constrained linear systems with parametric uncertainties. *Automatica* **117**, 1–8 (2020)
16. Çetin, M., Bahtiyar, B., Beyhan, S.: Adaptive uncertainty compensation-based nonlinear model predictive control with real-time applications. *Neural Comput. Appl.* **31**(2), 1029–1043 (2019)
17. Iplikci, S.: Runge-Kutta model-based adaptive predictive control mechanism for non-linear processes. *Trans. Inst. Meas. Control.* **35**(2), 166–180 (2013)
18. Aswani, A., Gonzalez, H., Sastry, S.S., et al.: Provably safe and robust learning-based model predictive control. *Automatica* **49**, 1216–1226 (2013)
19. Aswani, A., Bouffard, P., Zhang, X., et al.: Practical comparison of optimization algorithms for learning-based MPC with linear models. *arXiv preprint arXiv: 1404.2843* (2014)
20. Aswani, A., Taneja, J., Culler, D., et al.: Reducing transient and steady state electricity consumption in HVAC using learning-based model predictive control. *Proc. IEEE* **100**(1), 240–253 (2012)
21. Bouffard, P., Aswani, A., Tomlin, C.: Learning-based model predictive control on a quadrotor: onboard implementation and experimental results. *IEEE International Conference on Robotics and Automation*. Saint Paul, 279–284 (2012)
22. Aswani, A., Bouffard, P., Tomlin, C.: Extensions of learning-based model predictive control for real-time application to a quadrotor helicopter. *American Control Conference*, Montreal, 4661–4666 (2012)
23. Ioannou, P.A., Sun, J.: *Robust Adaptive Control*. Prentice-Hall, NJ, Upper Saddle River (1996)
24. Ibragimov, I.A., Hasminskii, R.Z.: *Statistical Estimation: Asymptotic Theory*. Springer, New York (1981)
25. Aranovskiy, S., Bobtsov, A., Ortega, R., et al.: Performance enhancement of parameter estimators via dynamic regressor extension and mixing. *IEEE Trans. Autom. Control* **62**(7), 3546–3550 (2017)
26. Chowdhary, G., Yucelen, T., Mühlegg, M., et al.: Concurrent learning adaptive control of linear systems with exponentially convergent bounds. *Int. J. Adapt. Control Signal Process.* **27**(4), 280–301 (2013)
27. Brunner, F.D., Heemels, W.P.M.H., Allgower, F.: Robust event-triggered MPC with guaranteed asymptotic bound and average sampling rate. *IEEE Trans. Autom. Control* **62**(11), 5694–5709 (2017)
28. Duan, Z.Y., Yan, H.S., Zheng, X.Y.: Robust model predictive control based on recurrent multi-dimensional Taylor network for discrete-time non-linear time-delay systems. *IET Control Theory Appl.* **14**(13), 1806–1818 (2020)
29. Limon, D., Alvarado, I., Alamo, T., et al.: Robust tube-based MPC for tracking of constrained linear systems with additive disturbances. *J. Process Control* **20**(3), 248–260 (2010)
30. Mayne, D.Q., Rawlings, J.B., Rao, C.V., et al.: Constrained model predictive control: stability and optimality. *Automatica* **36**(6), 789–814 (2000)
31. Chisci, J., Rossiter, A., Zappa, G.: Systems with persistent disturbances: predictive control with restricted constraints. *Automatica* **37**(7), 1019–1028 (2001)
32. Gilbert, E.G., Tan, K.T.: Linear systems with state and control constraints: the theory and application of maximal output admissible sets. *IEEE Trans. Autom. Control* **36**(9), 1008–1020 (1991)
33. Chenand, H., Allgower, F.: A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability. *Automatica* **34**(10), 1205–1217 (1998)
34. Fleming, J., Kouvaritakis, B., Cannon, M.: Robust tube MPC for linear systems with multiplicative uncertainty. *IEEE Trans. Autom. Control* **60**(4), 1087–1092 (2015)
35. Rakovic, S.V., Kouvaritakis, B., Cannon, M., et al.: Parameterized tube model predictive control. *IEEE Trans. Autom. Control* **57**(11), 2746–2761 (2012)
36. Rakovic, S.V., Baric, M.: Parameterized robust control invariant sets for linear systems: theoretical advances and computational remarks. *IEEE Trans. Autom. Control* **55**(7), 1599–1614 (2010)
37. Yan, Z., Le, X., Wang, J.: Tube-based robust model predictive control of nonlinear systems via collective neurodynamic optimization. *IEEE Trans. Industr. Electron.* **63**(7), 4377–4386 (2016)
38. Bemporad, A., Morari, M., Dua, V., et al.: The explicit linear quadratic regulator for constrained systems. *Automatica* **38**(1), 3–20 (2002)
39. Grimm, G., Messina, M.J., Tuna, S.E., et al.: Examples when nonlinear model predictive control is nonrobust. *Automatica* **40**(10), 1729–1738 (2004)
40. Johnson, C., Anderson, B.: Sufficient excitation and stable reduced-order adaptive IIR filtering. *IEEE Trans. Acoust. Speech Signal Process.* **29**(6), 1212–1215 (1981)
41. Dong, H., Hu, Q., Akella, M.R., et al.: Composite adaptive attitude tracking control with parameter convergence under

- finite excitation. *IEEE Trans. Control Syst. Technol.* (2019). <http://doi:https://doi.org/10.1109/tcst.2019.2942802>
42. Mayne, D.Q., Falugi, P.: Stabilizing conditions for model predictive control. *Int. J. Robust Nonlinear Control* **29**(4), 894–903 (2019)
  43. Rawlings, J.B., Muske, K.R.: The stability of constrained receding horizon control. *IEEE Trans. Autom. Control* **38**(10), 1512–1516 (1993)
  44. Rawlings, J.B., Mayne, D.Q., Diehl, M.: *Model predictive control: theory, computation, and design*, 2nd ed. Madison. Nob Hill Publishing, 218–255 (2017)
  45. Mayne, D.Q., Kerrigan, E.C., Van Wyk, E.J., et al.: Tube-based robust nonlinear model predictive control. *Int. J. Robust Nonlinear Control* **21**(11), 1341–1353 (2011)
  46. Falugi, P., Mayne, D.Q.: Getting robustness against unstructured uncertainty: a tube-based MPC approach. *IEEE Trans. Autom. Control* **59**(5), 1290–1295 (2014)
  47. Pradhan, S.K., Subudhi, B.: Nonlinear adaptive model predictive controller for a flexible manipulator: an experimental study. *IEEE Trans. Control Syst. Technol.* **22**(5), 1754–1768 (2014)
  48. Fisac, J.F., Akametalu, A.K., Zeilinger, M.N., et al.: A general safety framework for learning-based control in uncertain robotic systems. *IEEE Trans. Autom. Control* **64**(7), 2737–2752 (2019)
  49. Zhang, H., Liu, D., Luo, Y., et al.: *Adaptive dynamic programming for control: algorithms and stability*. Springer, Berlin (2012)
  50. Mosek. Accessed: Apr (2017). <https://www.mosek.com/downloads>
  51. Kali, Y., Saad, M., Benjelloun, K.: Discrete-time second order sliding mode with time delay control for uncertain robot manipulators. *Robot. Auton. Syst.* **94**, 53–60 (2017)
  52. Kolmanovsky, I., Gilbert, E.: Theory and computation of disturbance invariant sets for discrete-time linear systems. *Math. Probl. Eng.* **4**, 317–367 (1998)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.