Received: 13 March 2020 | Revised: 10 June 2020 | Accepted: 7 September 2020

IET Control Theory & Applications

DOI: 10.1049/cth2.12039

ORIGINAL RESEARCH PAPER

Linearized Bregman iteration based model-free adaptive sliding mode control for a class of non-linear systems

Shouli Gao[1] | Dongya Zhao[1] | Xinggang Yan[2] | Sarah K. Spurgeon[3]

[1] College of New Energy, China University of Petroleum (East China), Qingdao, China

[2] School of Engineering and Digital Arts, University of Kent, Canterbury, UK

[3] Department of Electronic & Electrical Engineering, University College London, Torrington Place, London, UK

Correspondence
Dongya Zhao, College of New Energy, China University of Petroleum (East China), Qingdao, China, 266580.
Email: dyzhao@upc.edu.cn

Funding information
The National Nature Science Foundation of China, Grant/Award Number: 61973315

Abstract

There is a growing demand for robust data-driven control methods particularly for industrial process control. This paper presents a new model-free adaptive sliding mode control approach for a class of discrete-time, multiple input and multiple output non-linear systems. The proposed methodology seeks to address issues with the computation of inverse matrices and problems with singularity in existing methods while at the same time seeking to enhance robustness. A Majorization–Minimization technique and the L_1 norm are used within the proposed optimization and an online iterative approach is described for update of the control law. The closed-loop system response is proved to be stable. The effectiveness of the proposed control is validated by extensive simulation and also experimental results, with the performance obtained by the proposed approach being compared throughout with a well-known approach from the established literature.

1 | INTRODUCTION

In modern control theory, the design of the controller is frequently carried out using a mathematical model. Such methods have been widely used and have achieved remarkable results in many industrial processes. However, in some cases, it may be difficult or unduly time-consuming to construct accurate models. At the same time, a considerable number of process industries generate and store large amounts of data every day, which may be very valuable to underpin a data-based control approach. For this reason, increasing attention is being paid to the development and application of data-driven control methods [1].

Data-driven control can be divided into two categories: offline data-driven control [2–4] and online data-driven control [5–7]. Offline data-driven control is suitable for offline large-scale data processing and is not amenable to complex iterative calculations. Online data-driven control can estimate system models in real time. Among these methods, iterative learning control has been found to provide high precision with repetitive tasks although the same initial point and desired trajectory may be required in each phase of operation [8]. Model-free adaptive control (MFAC) can estimate a system model online without the limitations required by iterative learning control [9, 10]. However, the pseudo-partial derivative (PPD) matrix in MFAC is slowly time-varying, which makes it difficult to accommodate sudden disturbances while maintaining high levels of performance.

Due to its established advantages which include robustness to parameter uncertainties and external disturbances, rapid response and straightforward implementation, sliding mode control (SMC) has received extensive attention in the literature and has been successfully applied to various complex systems [11–13]. The study of discrete-time SMC has facilitated the integration of sliding mode techniques with industrial computer systems [14, 15]. However, these studies mainly focus on the development of adaptive model-based controllers. The design of a model-free approach to adaptive sliding mode control has to date received less attention in the literature. A discrete-time control approach makes it easier to integrate the SMC paradigm with data-driven approaches. Hence, the combination of MFAC with SMC has become an attractive proposition to deal with non-linear discrete-time systems where the model is unknown and external disturbances are present [16].

There have been some results obtained by combining MFAC with SMC [16–18]. Model-free adaptive sliding mode control (MFASMC) is proposed in [16] for single input and single output non-linear discrete systems. MFASMC for the specific case of robot dynamics is proposed in [17]; this extends the control approach from single input single output to multiple input multiple output (MIMO). An extended state observer is incorporated into the controller design to construct a model-free adaptive sliding surface which enables decoupling control of the system in [18]. However, the effect of a reset algorithm which is required in practical control implementation to prevent singularity of the control has not been considered by the mentioned methods. Meanwhile, the L_2 norm is typically used in the control input criterion function and a parameter matrix inversion appears in the control algorithm. In order to avoid a matrix inverse in the applied control, the norm of the matrix has been used to replace the original matrix in the implemented control. This overcomes potential problems with singularity but may, however, degrade the control performance [16–18]. It should be noted that the same matrix inversion exists in the computation of the sliding mode reaching law, which adds to the controller online computation time.

The purpose of this study is to design a novel MFASMC for a class of MIMO nonlinear discrete-time systems where the model is unknown based on linear Bregman iteration [19]. The key point of this study is to use linear Bregman iteration and Majorization–Minimization theory [20] to facilitate controller design. Unlike the conventional MFAC-based SMC strategies [16–18], the proposed MFASMC has strong robustness while negating the need for matrix inverse calculations. The main contributions of this study are summarized as follows: (i) the Majorization–Minimization (MM) principle is used in optimizing the objective function, which avoids the computational time increase caused by matrix inversion; (ii) an L_1 norm term is used in the control input criterion function and the linear Bregman algorithm is employed for the solution, which can enhance the system robustness.

The paper is organized as follows: Section 2 formulates the control problem and some assumptions are provided; Section 3 designs the novel MFASMC; the corresponding stability analysis is given in Section 4; Section 5 gives simulation and experimental results; Some conclusions are given in Section 6.

2 | PROBLEM FORMULATION

Consider a MIMO non-linear discrete-time system described by

$$y(k+1) = F(y(k), \dots, y(k-n_y), \quad u(k), \dots, u(k-n_u)), \tag{1}$$

where $u(k) = [u_1(k), u_2(k), \dots, u_n(k)]^T \in R^n$ and $y(k) = [y_1(k), y_2(k), \dots, y_n(k)]^T \in R^n$ are input and output vectors, respectively. $F(y(k), \dots, y(k-n_y), u(k), \dots, u(k-n_u)) \in R^n$ is an unknown non-linear function. $n_y, n_u > 0$ are unknown system orders.

For system (1), the following assumptions are given [9]:

Assumption 1. *The partial derivatives of $F(y(k), \dots, y(k-n_y), u(k), \dots, u(k-n_u))$ with respect to control inputs $u(k)$ are continuous and non-zero.*

Assumption 2. *System (1) satisfies generalized Lipschitz condition, that is, $\|\Delta y(k+1)\| < b\|\Delta u(k)\|$ for any k and $\|\Delta u(k)\| \neq 0$, where $b > 0$, $\Delta y(k+1) = y(k+1) - y(k)$, $\Delta u(k) = u(k) - u(k-1)$.*

Assumption 3. *The function $F(y(k), \dots, y(k-n_y), u(k), \dots, u(k-n_u))$ in (1) is locally continuous and bounded.*

Remark 1. From a practical point of view, these assumptions are reasonable. Assumption 1 is a typical condition for many control designs and is satisfied by many non-linear systems. Assumption 2 poses a limitation on the rate of change of the system output in response to an applied control input. Assumption 3 is appropriate for many non-linear systems.

Two Lemmas from [9] are now presented.

Lemma 1. *For the MIMO non-linear discrete-time system (1) satisfying Assumptions 1 and 2, there exists a $\Phi(k)$, which is called the PPD matrix, such that system (1) can be written as*

$$\Delta y(k+1) = \Phi^T(k)\Delta u(k) \tag{2}$$

for bounded $\Phi(k) = [\Phi(k_{ij})]_{n\times n}$.

Remark 2. This type of data-driven control does not need explicit use of a mathematical model for the controller analysis and design. The design procedure does not require the model uncertainty to be directly characterized.

Lemma 2. *Consider system (2). If Assumptions 1 and 2 are satisfied, the unknown matrix $\Phi(k)$ can be determined from the following data-driven adaptive law:*

$$\hat{\Phi}(k) = \hat{\Phi}(k-1) + \frac{\chi(\Delta y(k-1) - \hat{\Phi}(k-1)\Delta u(k-1))}{\tau + \|\Delta u(k-1)\|^2} \Delta u^T(k-1), \tag{3}$$

where $\tau > 0$, and $\chi > 0$ are weighting factors and $\hat{\Phi}(k) = [\hat{\Phi}(k_{ij})]_{n\times n}$ is an estimate of the unknown $\Phi(k)$.

Remark 3. It is assumed that $\Phi(k)$ is a diagonally dominant matrix in the following sense: $|\Phi_{ij}(k)| \le b_1$, $b_2 \le |\Phi_{ii}(k)| \le ab_2$, $a \ge 1$, $b_2 > b_1(2a+1)(m-1)$, $i = 1, \dots, n$, $j = 1, \dots, n$, $i \neq j$. The signs of the elements of $\Phi(k)$ are fixed as in [10, 17, 18]. To guarantee $\hat{\Phi}$ is non-singular, (3) is modified to include the following reset algorithm [10]:

$$\hat{\Phi}_{ii}(k) = \hat{\Phi}_{ii}(0), \quad if\,|\hat{\Phi}_{ii}(k)| > b_2 \quad or\ |\hat{\Phi}_{ii}(k)| > ab_2$$
$$or\ sign(\hat{\Phi}_{ii}(k)) \neq sign(\hat{\Phi}_{ii}(0)),$$
$$\hat{\Phi}_{ij}(k) = \hat{\Phi}_{ij}(0), \quad if\,|\hat{\Phi}_{ij}(k)| > b_1$$
$$or\ sign(\hat{\Phi}_{ij}(k)) \neq sign(\hat{\Phi}_{ij}(0)), \quad i \neq j,$$

TABLE 1 Controller parameters

MFASMC-L1	$\tau = 0.1, \chi = 0.1, \bar{\lambda} = 0.1, \bar{\gamma}_i = 0.1, q = 1, \kappa = 1, T = 0.01$
MFASMC-L2	$\tau = 0.1, \chi = 0.1, \lambda = 0.1, \gamma_i = 0.1, q = 1, \kappa = 1, T = 0.01, \rho = 1, a = 4.5, b_1 = 0.01, b_2 = 0.1$

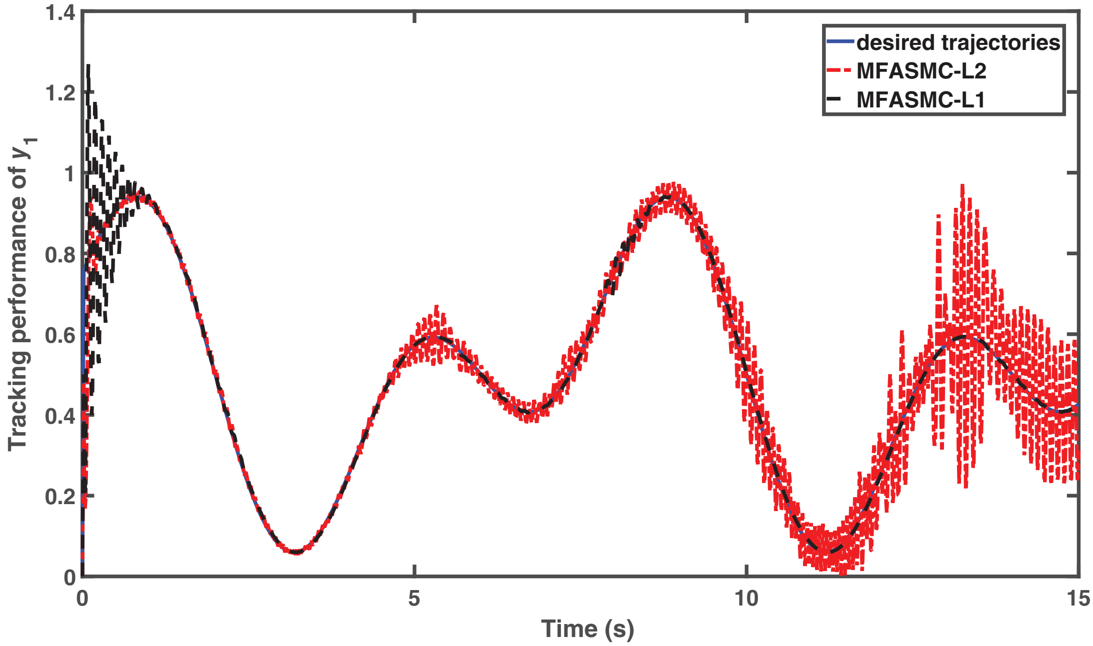

FIGURE 1 Tracking performance of y_1

where $\hat{\Phi}_{ij}(0)$ is the initial value of $\hat{\Phi}_{ij}(k)$, $i = 1, \ldots, n$ and $j = 1, \ldots, n$. Note that the reset algorithm will increase the online computation time and affect the control performance.

In order to make the system output track the desired trajectory and to achieve a smooth control input, an MAFC is designed using the following objective function [9]:

$$J(u(k)) = \|y_d(k+1) - y(k+1)\|_2^2 + \bar{\lambda}\|u(k) - u(k-1)\|_2^2, \tag{4}$$

where $y_d(k+1)$ is the desired trajectory and $\bar{\lambda} > 0$ is a weighting factor.

According to Lemmas 1 and 2, Equations (2) and (3), the following MAFC can be designed [9]:

$$u(k) = u(k-1) + (\bar{\lambda}I + \hat{\Phi}^T(k)\hat{\Phi}(k))^{-1}\hat{\Phi}^T(k)\Delta y_d(k), \tag{5}$$

where $\Delta y_d(k) = y_d(k+1) - y(k)$ and $\hat{\Phi}(k)$ is the adaptive estimate of $\Phi(k)$ given in (3).

TABLE 2 ITSE for the two methods

ITSE	y_1	y_2
MFASMC-L1	0.0496	0.1429
MFASMC-L2	0.3256	0.1430

The matrix inversion $(\bar{\lambda}I + \hat{\Phi}^T(k)\hat{\Phi}(k))^{-1}$ in Equation (5) will be computationally intensive when the dimensions of the system input and output are large. An alternative MAFC is designed in [10], whereby $\bar{\lambda}I + \Phi^T(k)\Phi(k)$ is replaced with $\bar{\lambda} + \|\hat{\Phi}(k)\|^2$ to yield:

$$u(k) = u(k-1) + \Delta\bar{u}_{\mathrm{MFA}}(k)$$

$$\Delta\bar{u}_{\mathrm{MFA}}(k) = \frac{\rho\hat{\Phi}^T(k)\Delta y_d(k)}{\bar{\lambda} + \|\hat{\Phi}(k)\|^2}, \tag{6}$$

where $\rho > 0$ is a weighting factor.

Remark 4. Note that the control (5) includes the matrix inversion $(\bar{\lambda}I + \hat{\Phi}^T(k)\hat{\Phi}(k))^{-1}$, which will increase the calculation time. On the other hand, the control (6) uses $\bar{\lambda} + \|\hat{\Phi}(k)\|^2$ to avoid the computation of the matrix inversion but this may be at the expense of control performance. Furthermore, the term $\|u(k) - u(k-1)\|_2^2$ is used in the objective function. It is well known that the L_2 norm is sensitive to external disturbances and system uncertainties, which can also reduce the robustness of the system [21].

Motivated from the above discussion, a data-driven MFASMC is developed for system 2 in [17, 18]. A conventional sliding surface is defined in the following linear form:

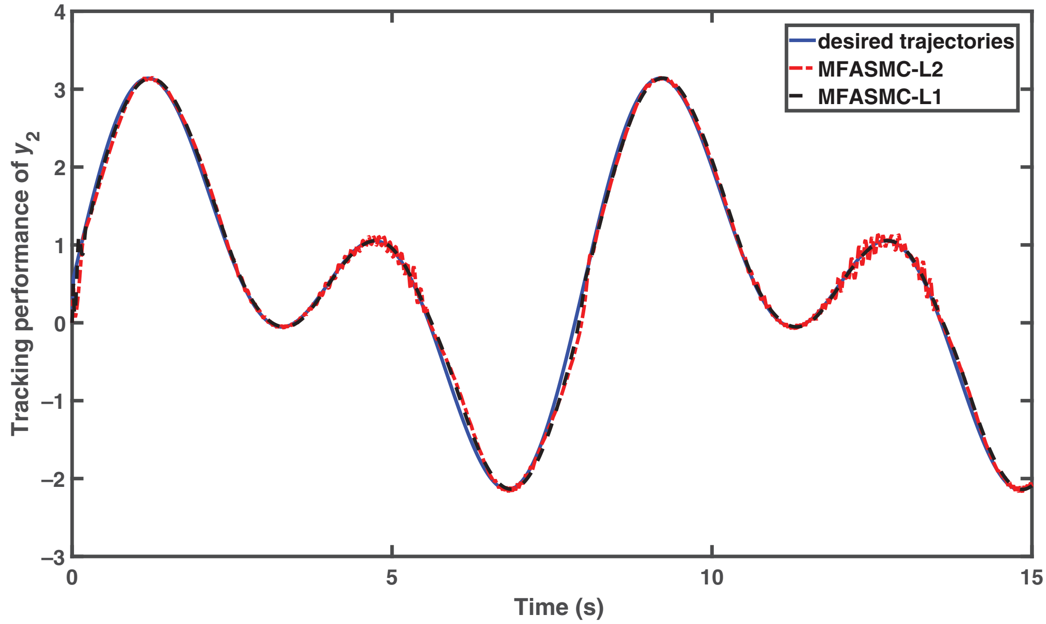

FIGURE 2 Tracking performance of y_2

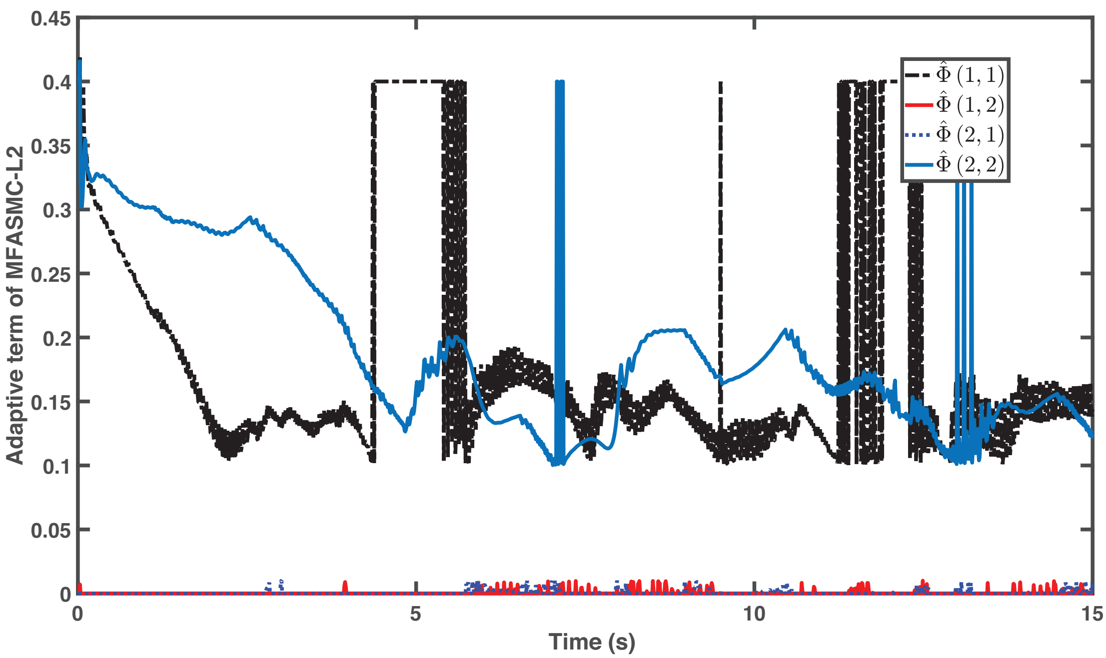

FIGURE 3 Adaptive term of MFASMC-L2

$$S(k) = e(k), \tag{7}$$

where $e(k) = y_d(k) - y(k) = [e_1(k), e_2(k), \dots, e_n(k)]^T \in R^n$. The reaching law is designed as follows:

$$S(k+1) - S(k) = -qTS(k) - \bar{\kappa}T\,\mathrm{sign}(S(k)), \tag{8}$$

where $q > 0$, $1 - qT > 0$, $\bar{\kappa} > 0$ and T is the discrete-time sample period. The reachability condition for a discrete sliding mode may be expressed as $\|s(k+1)\| < \|s(k)\|$ [22]. This guarantees that from any initial state the discrete system will reach the sliding surface.

Now consider the system (2) where Assumptions 1, 2 and 3 are satisfied. A sliding surface (7) has been designed for system (2) and the unknown matrix $\Phi(k)$ is estimated by the adaptive law (3). The following data-driven MFASMC is defined in [17] as follows:

$$\begin{aligned} u(k) &= u(k-1) + \Delta\bar{u}(k), \\ \Delta\bar{u}(k) &= \Delta\bar{u}_{\mathrm{MFA}}(k) + \bar{\Gamma}\Delta\bar{u}_{\mathrm{SM}}(k), \end{aligned} \tag{9}$$

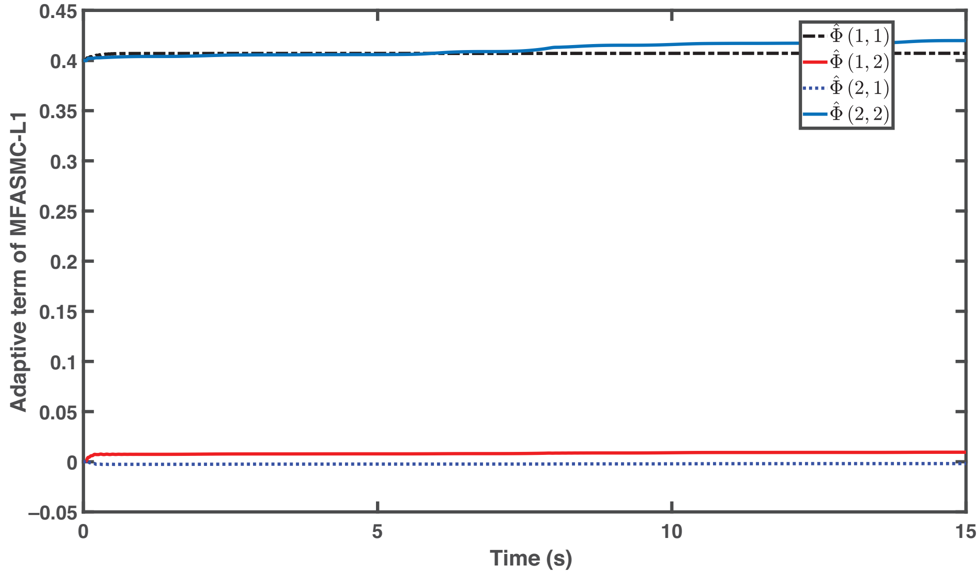

FIGURE 4 Adaptive term of MFASMC-L1

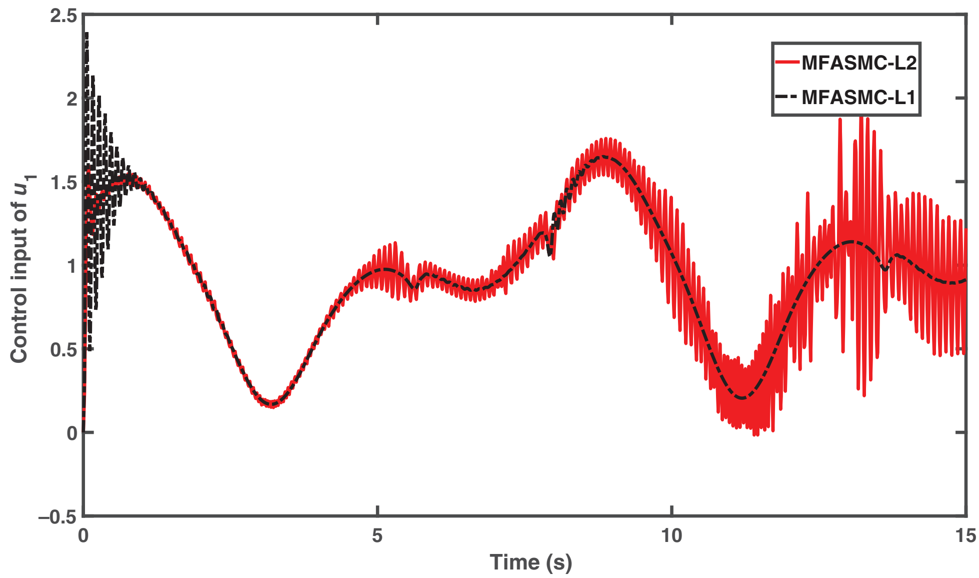

FIGURE 5 Control input of u_1

where $\Delta\bar{u}_{\rm MFA}(k)$ is defined in Equation (6), $\bar{\Gamma} = \mathrm{diag}([\bar{\gamma}_1, \ldots, \bar{\gamma}_n]) \in R^{n\times n}$ with $0 < \bar{\gamma}_i \le 1$ $(i = 0, 1, \ldots, n)$ weighting factors and

$$\Delta\bar{u}_{\rm SM}(k) = \hat{\Phi}^{-1}(k)(y_d(k+1) - y_d(k) + qTS(k) + \bar{\kappa}T\,\mathrm{sign}(S(k))). \tag{10}$$

Remark 5. Note that the inverse $\hat{\Phi}^{-1}(k)$ is still required in the control computation (9)–(10), which will contribute to the computational load. This paper proposed overcome the disadvantage in [17, 18].

The contribution of this paper can be summarized as follows. A new MFASMC is proposed for a class of MIMO non-linear systems. System robustness is enhanced by using the MM principle and the L_1 norm in the design of the MFASMC. Furthermore, matrix inversion is avoided in the new control, which will effectively decrease the online computation time.

3 | MFASMC DESIGN

Inspired by the fundamental results on Linearized Bregman iteration and MM in [19, 20], respectively, the proposed control scheme is designed as follows:

$$u(k) = u(k-1) + \Delta u(k),$$

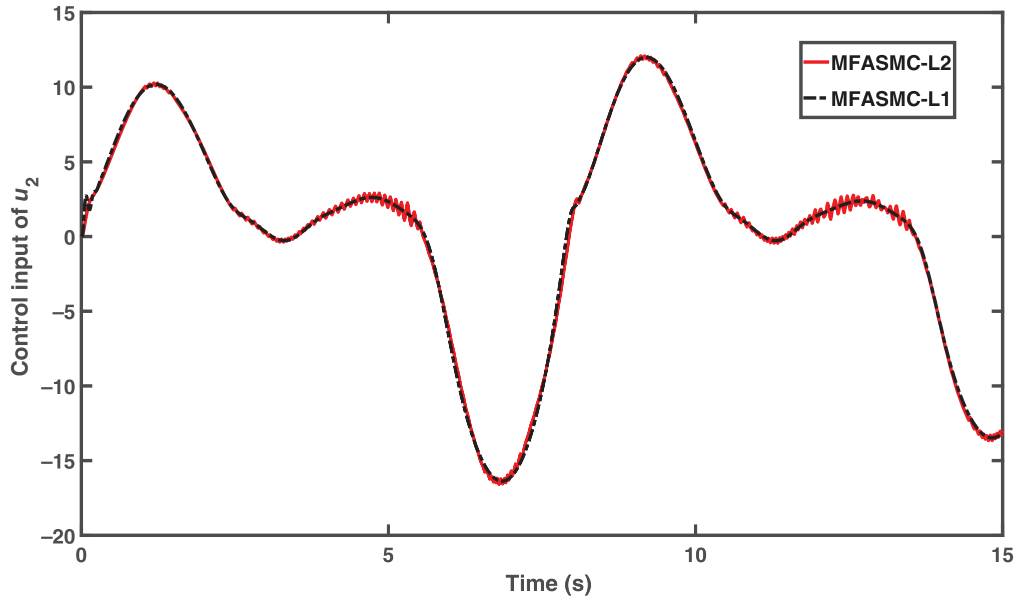

FIGURE 6 Control input of u_2

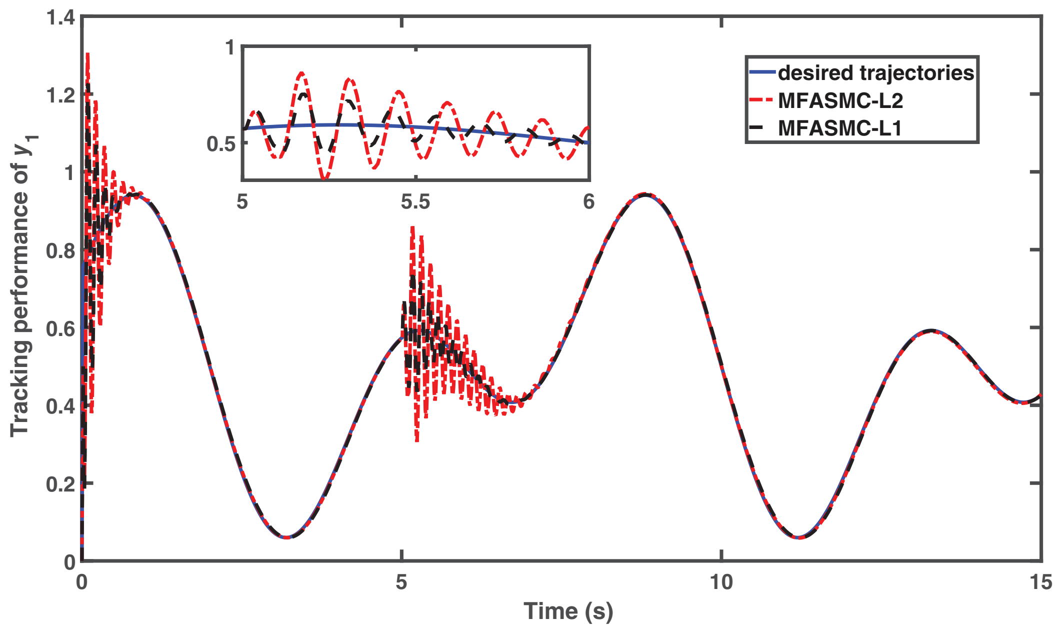

FIGURE 7 Tracking performance of y_1

$$\Delta u(k) = \Delta u_{MFA}(k) + \Gamma \Delta u_{SM}(k), \tag{11}$$

where $\Gamma = \text{diag}([\gamma_1, \dots, \gamma_n]) \in R^{n \times n}$, $0 < \gamma_i \leq 1$ $(i = 1, 2, \dots, n)$ are weighting factors, $\Delta u_{MFA}(k) \in R^n$ is the equivalent control law and $\Delta u_{SM}(k) \in R^n$ is a discontinuous element of the control law. The components $\Delta u_{MFA}(k)$ and $\Delta u_{SM}(k)$ are obtained by using Algorithms 1 and 2, respectively.

Remark 6. Compared with continuous-time system, the study on discrete-time system are very difficult. Specifically, the associated results of MFASMC for non-linear discrete systems are very rare. This paper designs a novel MFASMC for a class of MIMO nonlinear discrete-time systems. When compared with the control law proposed in [17], the above approach replaces the matrix inversion in (5) with iterative calculations and replaces the L_2 norm with the L_1 norm in the formulation of the objective function. The avoidance of inverse calculation enhances the efficiency of online calculation of the controller. It is noteworthy that the L_2 norm variance is sensitive to noises with large norms [21, 23]. So substituting the L_1 norm variance for the L_2 norm in the objective function is a meliorative strategy. This method can enhance the closed-loop system

Algorithm 1

Step 1: Establish the objective function relating to the equivalent control law:

$$G_{MFA}(\Delta u(k)) = \alpha \left\| \vartheta(k) + \frac{\hat{\Phi}^T(k)}{\alpha}(\Delta y_d(k) - \hat{\Phi}(k)\vartheta(k)) - \Delta u(k) \right\|_2^2 \quad (12)$$
$$+\lambda\|\Delta u(k)\|_1,$$

where $\alpha > eig_{\max}$, $eig_{\max}$ is the maximum eigenvalue of $\hat{\Phi}^T(k)\hat{\Phi}(k)$, $0 < \lambda < 1$ is weighting constant and $\vartheta(k) \in R^n$ is an instrumental variable.

Step 2: Solve $\vartheta(k)$ by using the following iterative calculation:

$$\vartheta_\varrho(k) = \vartheta_{\varrho-1}(k) + \frac{\hat{\Phi}^T(k)}{\alpha}\left(\Delta y_d(k) - \hat{\Phi}(k)\vartheta_{\varrho-1}(k)\right), \quad (13)$$

where $\vartheta_\varrho(k) \in R^n$ is an iterative loop variable, $\varrho = 1, 2, \ldots$ is the iteration number and the initial value of $\vartheta_0(k)$ is an arbitrary vector.

Step 3: If $\frac{\|\vartheta_\varrho(k)-\vartheta_{\varrho-1}(k)\|}{\|\vartheta_\varrho(k)\|} < \varepsilon$ then stop the iteration and let $\vartheta(k) = \vartheta_\varrho(k)$. Otherwise, let $\varrho = \varrho + 1$ and go to Step 2, where ε represents the precision of the calculation.

Step 4: Optimize (12) by using the linearized Bregman approach. The parameter update law for $\Delta u_{MFA}(k)$ is given as follows:

$$\begin{cases} U_{j+1}(k) = U_j(k) - \frac{2\alpha}{\lambda}\left(\Delta u_j(k) - \Delta u_{j-1}(k)\right) - \frac{\alpha}{\lambda}\left(\Delta u_{j-1}(k) - v(k)\right), \\ \Delta u_{j+1}(k) = 2\alpha T_\lambda\left(\hat{\Phi}^T(k)U_{j+1}(k)\right), \end{cases} \quad (14)$$

where $U_j(k) \in R^n$ and $\Delta u_j(k) \in R^n$ are iterative loop variables, $j = 1, 2, \ldots$ is an iteration number and the initial values of $U_0(k) \in R^n$ and $\Delta u_0(k) \in R^n$ are zero vectors, $v(k) = \vartheta(k) + \frac{\hat{\Phi}^T(k)}{\alpha}(\Delta y_d(k) - \hat{\Phi}(k)\vartheta(k))$, $T_\lambda(w) = [t_\lambda(w(1)), \ldots, t_\lambda(w(n))]^T$, and $t_\lambda(\cdot) \in R$ is given as

$$t_\lambda(\xi) = \begin{cases} 0, & if\, |\xi| \le \lambda \\ \text{sign}(\xi)(|\xi| - \lambda), & if\, |\xi| > \lambda \end{cases} \quad (15)$$

Step 5: If $\frac{\|\Delta u_j(k)-\Delta u_{j-1}(k)\|}{\|\Delta u_j(k)\|} < \delta$ then stop iteration and let $\Delta u_{MFA}(k) = \Delta u_j(k)$. Otherwise, let $j = j + 1$ and go to Step 4, where δ denotes the precision of the iteration.

End of Algorithm 1.

Algorithm 2

Step 1: Establish an objective function for the discontinuous control law:

$$G_{SM}(\Delta u(k)) = \left\| \vartheta^*(k) + \frac{\hat{\Phi}^T(k)}{\alpha}\right.$$
$$\left.\left(y_d(k+1) - y_d(k) + qTS(k) + \kappa T sign(S(k)) - \hat{\Phi}(k)\vartheta^*(k)\right) - \Delta u(k)\right\|_2^2, \quad (16)$$

where $\vartheta^*(k) \in R^n$ is an instrumental variable, $\kappa > 0$.

Step 2: Solve $\vartheta^*(k)$ by using the following iterative calculation:

$$\vartheta_l^*(k) = \vartheta_{l-1}^*(k) + \frac{\hat{\Phi}^T(k)}{\alpha}\left(\Delta y_d^*(k) - \hat{\Phi}(k)\vartheta_{l-1}^*(k)\right), \quad (17)$$

where $\vartheta_l^*(k) \in R^n$ is an iterative loop variable, $l = 1, 2, \ldots$ is an iteration number and the initial value of $\vartheta_0^*(k)$ is an arbitrary vector, $\Delta y_d^*(k) = y_d(k+1) - y_d(k) + qTS(k) + \kappa T sign(S(k))$.

Step 3: If $\frac{\|\vartheta_l^*(k)-\vartheta_{l-1}^*(k)\|}{\|\vartheta_l^*(k)\|} < \varepsilon^*$ then stop iteration and let $\vartheta^*(k) = \vartheta_l^*(k)$. Otherwise, let $l = l + 1$ and go to Step 2, where ε^* represents the precision of the iteration.

Step 4: Optimize (16) where the discontinuous control law $\Delta u_{SM}(k)$ is obtained as follows:

$$\Delta u_{SM}(k) = \vartheta^*(k) + \frac{\hat{\Phi}^T(k)}{\alpha}\left(\Delta y_d^*(k) - \hat{\Phi}(k)\vartheta^*(k)\right). \quad (18)$$

End of Algorithm 2.

robustness when compared with traditional MFASMC. Furthermore, the proposed approach replaces matrix inversion $\hat{\Phi}^{-1}(k)$ in (10) with iterative calculations, which can further provide stronger robustness than the approach in [17]. The proposed strategy uses SMC to deal with the system uncertainty and disturbances. When the system state is in the sliding mode, it can be expected to have excellent insensitivity to parameter variations and external disturbances.

4 | STABILITY ANALYSIS

Lemma 3. *Under Assumptions 1, 2 and 3, the sequences* $\{\vartheta_\varrho(k)\}$ *and* $\{\vartheta_l^*(k)\}$ *are absolutely convergent.*

Proof. Equation (13) can be written as

$$\begin{aligned} \vartheta_\varrho(k) &= \vartheta_{\varrho-1}(k) + \frac{\hat{\Phi}^T(k)}{\alpha}\left(\Delta y_d(k) - \hat{\Phi}(k)\vartheta_{\varrho-1}(k)\right) \\ &= \left(I - \frac{\hat{\Phi}^T(k)\hat{\Phi}(k)}{\alpha}\right)^\varrho \vartheta_0(k) \\ &\quad + \sum_{\varpi=1}^{\varrho}\left(I - \frac{\hat{\Phi}^T(k)\hat{\Phi}(k)}{\alpha}\right)^{\varpi-1}\frac{\hat{\Phi}^T(k)}{\alpha}\Delta y_d(k). \end{aligned}$$

Since $\alpha > \lambda_{\max}$, $0 < \|I - \frac{\hat{\Phi}^T(k)\hat{\Phi}(k)}{\alpha}\| < 1$, then $\lim_{\varrho\to\infty}(I - \frac{\hat{\Phi}^T(k)\hat{\Phi}(k)}{\alpha})^\varrho = \mathbf{0}_{n\times n}$. Because $\|I - \frac{\hat{\Phi}^T(k)\hat{\Phi}(k)}{\alpha}\| < 1$, $\sum_{\varpi=1}^{\varrho}(I - \frac{\hat{\Phi}^T(k)\hat{\Phi}(k)}{\alpha})^{\varpi-1}$ is absolutely convergent [24]. Hence, $\{\vartheta_\varrho(k)\}$ is absolutely convergent.

In a similar way, $\{\vartheta_l^*(k)\}$ can be proved to be absolutely convergent. □

Remark 7. In theory, ϱ and l are required to be infinite to guarantee $\{\vartheta_\varrho(k)\}$ and $\{\vartheta_l^*(k)\}$ are absolutely convergent. Note that they appear as a power term and with the availability of modern hardware, the numbers of iterations can be expected to be large enough within a sampling time to ensure that the computational results approach the theoretical values.

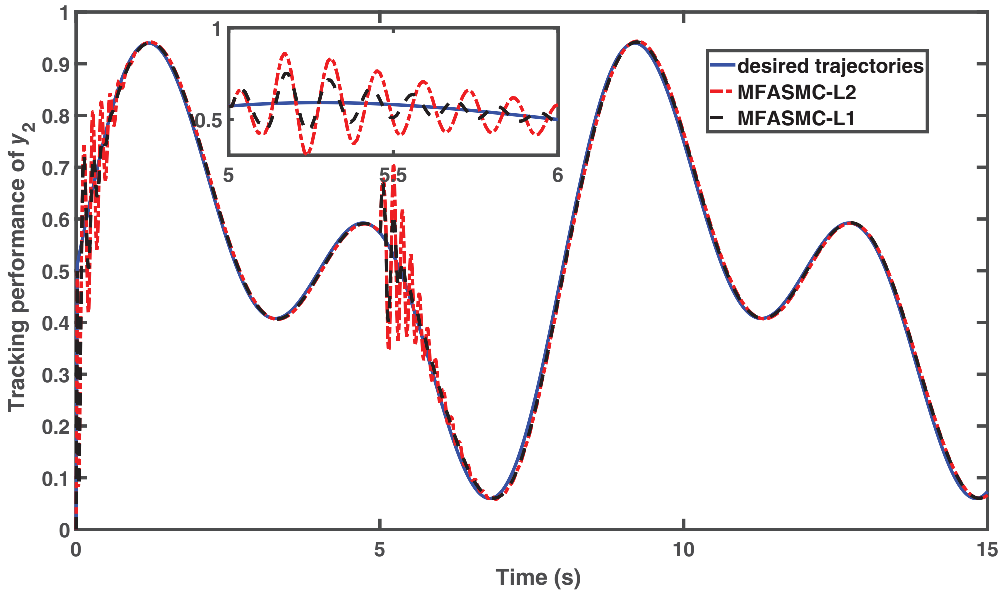

FIGURE 8 Tracking performance of y_2

Definition 1. *If H is a convex function, then the Bregman distance of H at two points A, B is defined as*

$$D_H^P(A,B) = H(A) - H(B) - \langle A - B, P\rangle, \tag{19}$$

where $P \in \partial H$ is a sub-gradient in the sub-differential of H at A red and $\langle A, B\rangle$ is the inner product of A and B[19].

Assumption 4. *The convex function $J(u)$ is continuously differentiable and there exists a positive constant β such that*

$$\|\partial J(u) - \partial J(v)\|^2 \leq \beta\langle \partial J(u) - \partial J(v), u - v\rangle, \forall u, v \in R^n, \tag{20}$$

where $\partial J(u)$ is the gradient of $J(u)$ [19].

Lemma 4. *Under Definition 1 and Assumption 4, if $\{\Delta u_j(k)\}$ is generated by the linearized Bregman iteration (14) with $\alpha > \lambda_{\max}$ and $\Delta u_j(k) \neq \Delta u_{j+1}(k)$, then*

$$\|\Delta u_j(k) - v(k)\| \leq \eta\|\Delta u_{j-1}(k) - v(k)\|, \tag{21}$$

where $0 < \eta < 1$.

Proof. See the Appendix. □

Theorem 1. *If $H(\Delta u(k))$ is convex and Assumption 4 is satisfied, then both of the sequences $\{\Delta u_j(k)\}$ and $\{U_j(k)\}$ are convergent and the convergence rate $\eta < 1$.*

Proof. Suppose that there exists an integer j^* such that $\Delta u_{j^*+1}(k) = \Delta u_{j^*}(k)$. From (14), it can be obtained that $\Delta u_j(k) = \Delta u_{j+1}(k)$ and $U_j(k) = U_{j+1}(k)$ for all $j > j^*$. Thus, both the sequences $\{\Delta u_j(k)\}$ and $\{U_j(k)\}$ converge.

Otherwise, suppose that $\Delta u_{j+1}(k) \neq \Delta u_j(k)$ for all j. Taking the inner product of both sides of (A.8) with respect to $\Delta u_j(k) - \Delta u_{j-1}(k)$

$$\begin{aligned}&\langle \lambda(U_{j+1} - U_j), \Delta u_j(k) - \Delta u_{j-1}(k)\rangle + 2\alpha\|\Delta u_j(k) - \Delta u_{j-1}(k)\|^2\\ &= -\alpha\langle(\Delta u_j(k) - v(k)), \Delta u_j(k) - \Delta u_{j-1}(k)\rangle.\end{aligned} \tag{22}$$

From (A.3)

$$\begin{aligned}&2\alpha\|\Delta u_{j+1}(k) - \Delta u_j(k)\|^2\\ &\leq -\alpha\langle(\Delta u_j(k) - v(k)), \Delta u_{j+1}(k) - \Delta u_j(k)\rangle\\ &\leq \alpha\|(\Delta u_j(k)) - v(k)\|\|\Delta u_{j+1}(k) - \Delta u_j(k)\|.\end{aligned}$$

Therefore,

$$\|\Delta u_{j+1}(k) - \Delta u_j(k)\| \leq \frac{1}{2}\|\Delta u_j(k) - v(k)\|. \tag{23}$$

Considering Lemma 4

$$\|\Delta u_j(k) - v(k)\| \leq \eta^j\|\Delta u_0(k) - v(k)\|. \tag{24}$$

Thus, for any $j^\star > j$

$$\begin{aligned}\|\Delta u_j(k) - \Delta u_{j^\star}(k)\| &\leq \sum_{\varsigma=j}^{j^\star}\|\Delta u_{\varsigma+1}(k) - \Delta u_\varsigma(k)\|\\ &\leq \frac{1}{2}\sum_{\varsigma=j}^{j^\star}\|v(k) - \Delta u_\varsigma(k)\|\end{aligned}$$

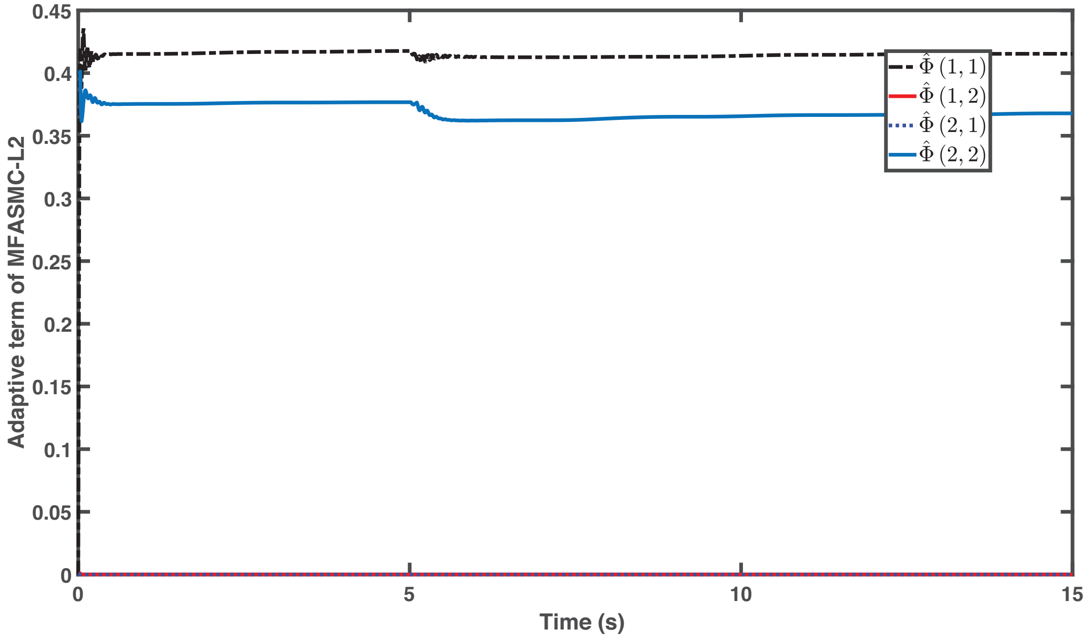

FIGURE 9 Adaptive term for MFASMC-L2

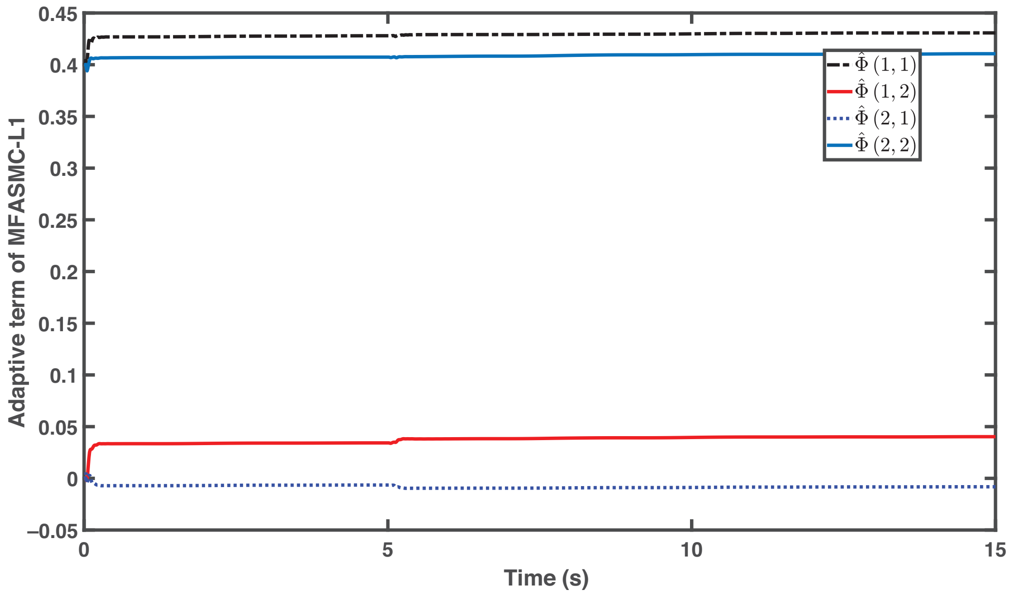

FIGURE 10 Adaptive term for MFASMC-L1

$$
\begin{aligned}
&\leq \frac{1}{2}\sum_{\varsigma=j}^{j^{\star}} \eta^{\varsigma}\|v(k)-\Delta u_0(k)\| \\
&\leq \frac{\eta^{j}}{2(1-\eta)}\|v(k)-\Delta u_0(k)\|.
\end{aligned}
$$

As $j^{\star} \to \infty$, it can be obtained that

$$
\left\|\Delta u_j(k)-\Delta u_{j^{\star}}(k)\right\| \leq \frac{\left\|\Delta u_0(k)-v(k)\right\|}{2(1-\eta)}\eta^{j}. \tag{25}
$$

The sequence $\{\Delta u_j(k)\}$ is a Cauchy sequence, which is convergent and the convergence rate is η. The convergence of $\{U_j(k)\}$ can be proved similarly. □

Theorem 2. *Consider system (1) where Assumptions 1–4 are satisfied and the control law is designed as (11). Then $S(k)$ and $e(k)$ are bounded.*

Proof. When the desired trajectory is constant, then

$$
\begin{aligned}
e(k+1) &= y_d - y(k+1) = y_d - y(k+1) - (y(k) - y(k)) \\
&= e(k) - \Phi^{T}(k)\Delta u(k).
\end{aligned} \tag{26}
$$

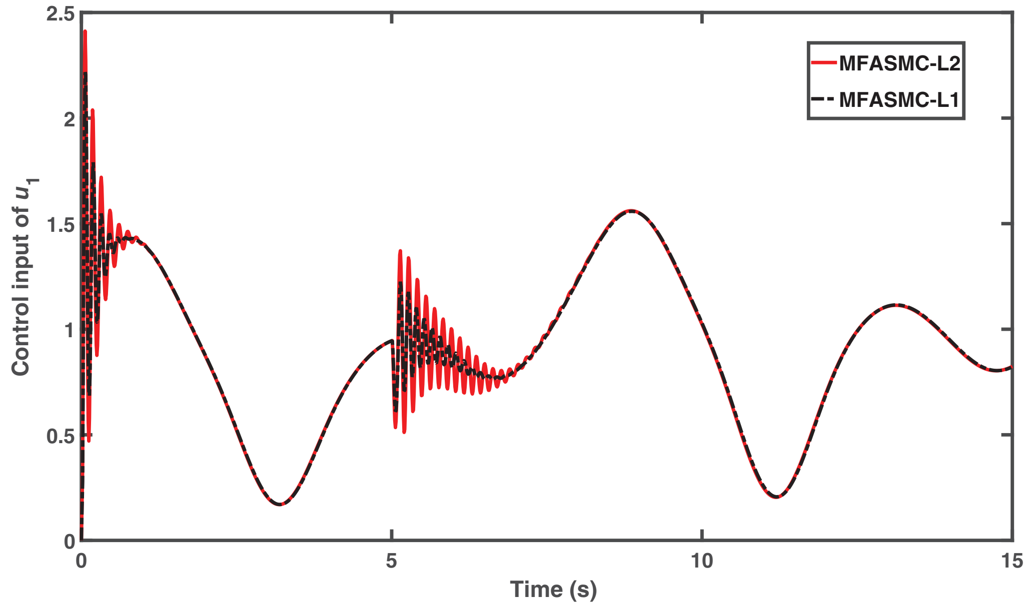

FIGURE 11 Control input u_1

Consider Lemmas 3 and 4, and substitute (11) into (26)

$$e(k+1) = e(k) - \frac{\hat{\Phi}^T(k)\hat{\Phi}(k)}{\alpha} e(k) - \Gamma qTS(k) - \Gamma \kappa T sign(S(k)). \tag{27}$$

Since $S(k) = e(k)$, then

$$\begin{aligned} S(k+1) - S(k) \\ = -\frac{\hat{\Phi}^T(k)\hat{\Phi}(k)}{\alpha} S(k) - \Gamma qTS(k) - \Gamma \kappa T sign(S(k)). \end{aligned} \tag{28}$$

It follows that

$$\begin{aligned} S(k) &= S(k-1) - \frac{\hat{\Phi}^T(k-1)\hat{\Phi}(k-1)}{\alpha} S(k-1) \\ &\quad - \Gamma qTS(k-1) - \Gamma \kappa T sign(S(k-1)) \\ &= \left(I - \frac{\hat{\Phi}^T(k-1)\hat{\Phi}(k-1)}{\alpha} - \Gamma qT \right) S(k-1) \\ &\quad - \Gamma \kappa T sign(S(k-1)) \\ &= \left(I - \frac{\hat{\Phi}^T(k-1)\hat{\Phi}(k-1)}{\alpha} - \Gamma qT \right)^k S(0) \\ &\quad - \sum_{\Upsilon=0}^{k-1} \left(I - \frac{\hat{\Phi}^T(k-1)\hat{\Phi}(k-1)}{\alpha} - \Gamma qT \right)^{k-1-\Upsilon} \\ &\quad \times \Gamma \kappa T sign(S(\Upsilon)). \end{aligned}$$

Since $\alpha > \lambda_{\max}$, $1 - qT > 0$ and $0 < \gamma_i \leq 1$, it follows that $0 < \| I - \frac{\hat{\Phi}^T(k)\hat{\Phi}(k)}{\alpha} - \Gamma qT \| < 1$. Then

$$\lim_{k \to \infty} \left(I - \frac{\hat{\Phi}^T(k-1)\hat{\Phi}(k-1)}{\alpha} - \Gamma qT \right)^k = \mathbf{0}_{n \times n}.$$

Meanwhile $\sum_{\Upsilon=0}^{k-1} (I - \frac{\hat{\Phi}^T(k-1)\hat{\Phi}(k-1)}{\alpha} - \Gamma qT)^{k-1-\Upsilon}$ is absolutely convergent when $k \to \infty$, so $S(k)$ is bounded. As $S(k) = e(k)$, then $e(k)$ is bounded.

When the desired trajectory is time-varying, the result can also be proved by the above method using the following augmented system to replace the system (1):

$$\begin{aligned} z(k+1) = F(y(k), \ldots, y(k-n_y), u(k), \ldots, u(k-n_u)) \\ - y_d(k+1). \end{aligned} \tag{29}$$

In a similar way, $S(k)$ and $e(k)$ can be proved to be bounded. $\square$

Remark 8. From Lemma 3, it can be seen that the Majorization–Minimization algorithm is convergent in Algorithms 1 and 2. The Majorization–Minimization algorithm converts the matrix inversion into iterative calculation, which increases the computational efficiency and avoids the impact of any reset of $\hat{\Phi}(k)$ on the controller performance. From Theorem 1, it can be seen that the objective function with the L_1 norm converges. Because of the disruptive effect of the L_2 norm variance at points with large norms, the L_1 norm is expected to be more robust to noises than the L_2 norm. From Theorem 2, it can be seen that $S(k)$ is bounded, which makes the system satisfy the existence and reachability conditions, ensuring the stability of

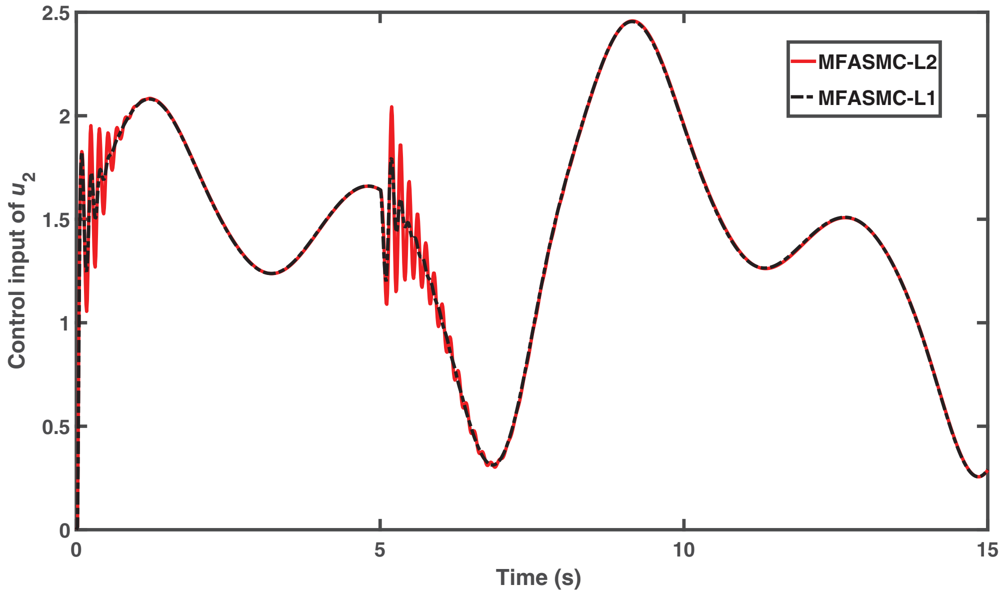

FIGURE 12 Control input u_2

the sliding mode. Therefore, the system state can be guaranteed to converge to the sliding surface, and it follows that the system will exhibit strong robustness to uncertainties.

5 | SIMULATION AND EXPERIMENTAL VERIFICATION

5.1 | Simulation results

In this section, the performance of the proposed control, which will be denoted MFASMC-L1, will be compared with that of the established MFASMC, denoted MFASMC-L2, from [17]. Two cases are analysed: Case 1 shows the influence of substituting the matrix inversion with the proposed iterative approach on system performance; Case 2 shows that using the L_1 norm improves system robustness.

The parameters of the two control laws are listed in Table 1. The iteration accuracy is given as $\varepsilon^* = 0.001$ and $\varepsilon = 0.001$. In MFASMC-L1 $\delta = 0.001$.

A MIMO non-linear system from [25] is used for the simulation testing, which satisfies the minimal controllability conditions. The dynamic model is given by

$$\begin{cases} x_{11}(k+1) = \dfrac{x_{11}^2(k)}{1+x_{11}^2(k)} + 0.3x_{12}(k) + d_1 \\ x_{12}(k+1) = \dfrac{x_{11}^2(k)}{1+x_{12}^2(k)+x_{21}^2(k)+x_{22}^2(k)} + A(k)u_1(k) \\ x_{21}(k+1) = \dfrac{x_{21}^2(k)}{1+x_{21}^2(k)} + 0.2x_{22}(k) + d_2 \\ x_{22}(k+1) = \dfrac{x_{21}^2(k)}{1+x_{11}^2(k)+x_{12}^2(k)+x_{22}^2(k)} + B(k)u_2(k), \end{cases}$$

where $A(k) = 1 + 0.1\sin(2\pi k/1500)$, $B(k) = 1 + 0.1\cos(2\pi k/1500)$. $y_1 = x_{11}$, $y_2 = x_{21}$ denote the system outputs. The initial conditions are given as $x_{11}(0) = x_{21}(0) = 0.5$, $x_{12}(0) = x_{22}(0) = 0$, $\hat{\Phi}(0) = \begin{bmatrix} 0.4 & 0 \\ 0 & 0.4 \end{bmatrix}$.

Case 1: Investigation of the influence of substituting the matrix inversion with the proposed iterative approach

The desired trajectory is given by

$$y_{d1}(k) = 0.5 + 0.25\cos(0.25\pi/100) + 0.25\sin(0.25\pi/100)$$

$$y_{d2}(k) = 0.5 + 1.5\cos(0.25\pi/100) + 1.5\sin(0.25\pi/100).$$

In this case, no disturbance is present so that

$$\begin{cases} d_1 = 0 \\ d_2 = 0. \end{cases}$$

Figures 1 and 2 show the tracking performance. It can be observed that MFASMC-L1 tracks the desired trajectories well, while the tracking accuracy with MFASMC-L2 is worse. Figure 3 shows the time evolution of the adaptive law for MFASMC-L2. When the adaptive law approaches the singular point, there is a reset in the adaptation. It can be clearly seen from the simulation results that when the adaptive law starts to reset, the tracking performance worsens. Figure 4 shows the performance of the adaptive law for MFASMC-L1. It can be seen that the response is bounded and smooth. The proposed MFASMC-L1 method uses MM technology to avoid matrix inversion. So it can achieve better tracking performance than that obtained using MFASMC-L2. Figures 5 and 6 show the control input of the two methods, respectively. Comparing the results shows that MFASMC-L2 exhibits large oscillations when compared to

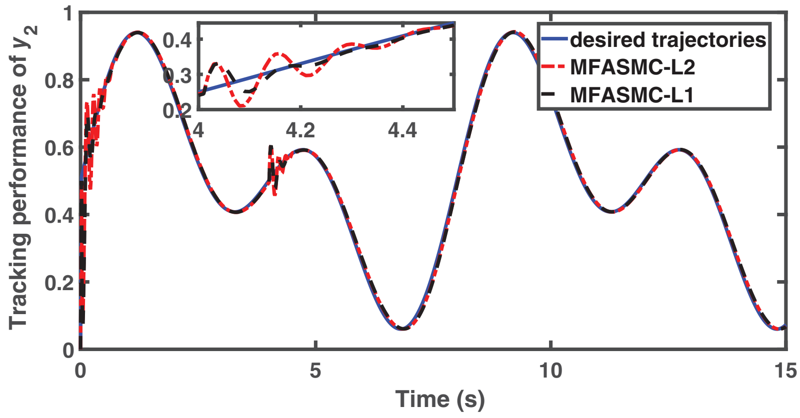

FIGURE 13 Tracking performance of y_1

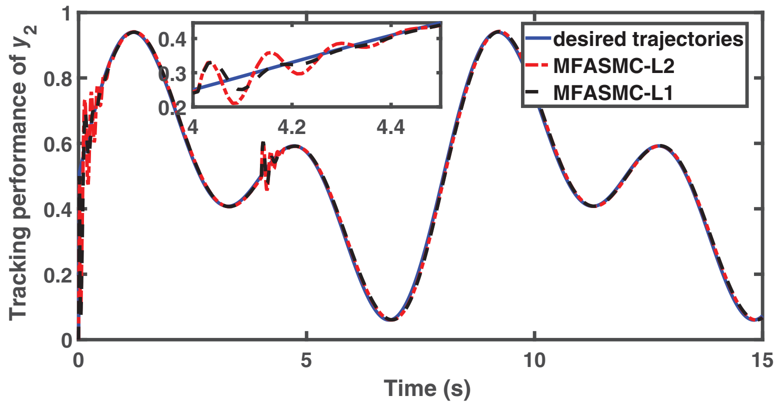

FIGURE 14 Tracking performance of y_2

the proposed approach. The Integral of Time Squared Error (ITSE) for the two control methods is listed in Table 2. It is clear that the ITSE of MFASMC-L1 is much smaller than that obtained for MFASMC-L2. The simulation results validate the effectiveness of the proposed approach. The iterative procedure proposed is able to overcome problems with matrix inversion and singularity that are present in existing methods (see, e.g. [16, 17]).

Case 2: The control performance in the presence of external disturbances

The desired trajectory is given by

$$y_{d1}(k) = 0.5 + 0.25\cos(0.25\pi/100) + 0.25\sin(0.25\pi/100)$$

$$y_{d2}(k) = 0.5 + 0.25\cos(0.25\pi/100) + 0.25\sin(0.25\pi/100).$$

The disturbance is

TABLE 3 ITSE for the two methods

	Disturbance (30)		Disturbance (31)	
ITSE	y_1	y_2	y_1	y_2
MFASMC-L1	0.0301	0.0186	0.0277	0.0172
MFASMC-L2	0.0394	0.0210	0.0301	0.0176

$$\begin{cases} d_1 = 0.05, d_2 = 0.05 & \text{if } 5 < T < 5.1, \\ d_1 = 0, d_2 = 0 & \text{otherwise.} \end{cases} \tag{30}$$

Figures 7 and 8 show the tracking control performance. By comparing the performance of the two methods, the proposed approach converges to the desired trajectory more quickly in the presence of a disturbance (30). Due to the use of the L_2 norm, the MFASMC-L2 is more sensitive to external disturbances. The sparsity of the L_1 norm can deal with disturbances more effectively, which renders the proposed approach more robust. Figures 9 and 10 show the performance of the adaptive law for the two methods. It can be seen that the adaptive law for MFASMC-L2 does not reset and both adaptive laws are bounded in the absence of singular effects. Figures 11 and 12 show the control inputs for both methods. By comparing Figures 11 and 12, the proposed method is observed to oscillate less when the disturbance occurs.

To further test the system robustness, the previous impulsive disturbance (30) is replaced by the following step disturbance (31):

$$\begin{cases} d_1 = 0.05, d_2 = 0.05 & \text{if } T > 4 \\ d_1 = 0, d_2 = 0 & \text{otherwise.} \end{cases} \tag{31}$$

Figures 13 and 14 show the tracking performance. It can be seen that due to the adopted sliding mode control, the system state in both cases converges to the desired trajectory in the presence

FIGURE 15 Experimental equipment

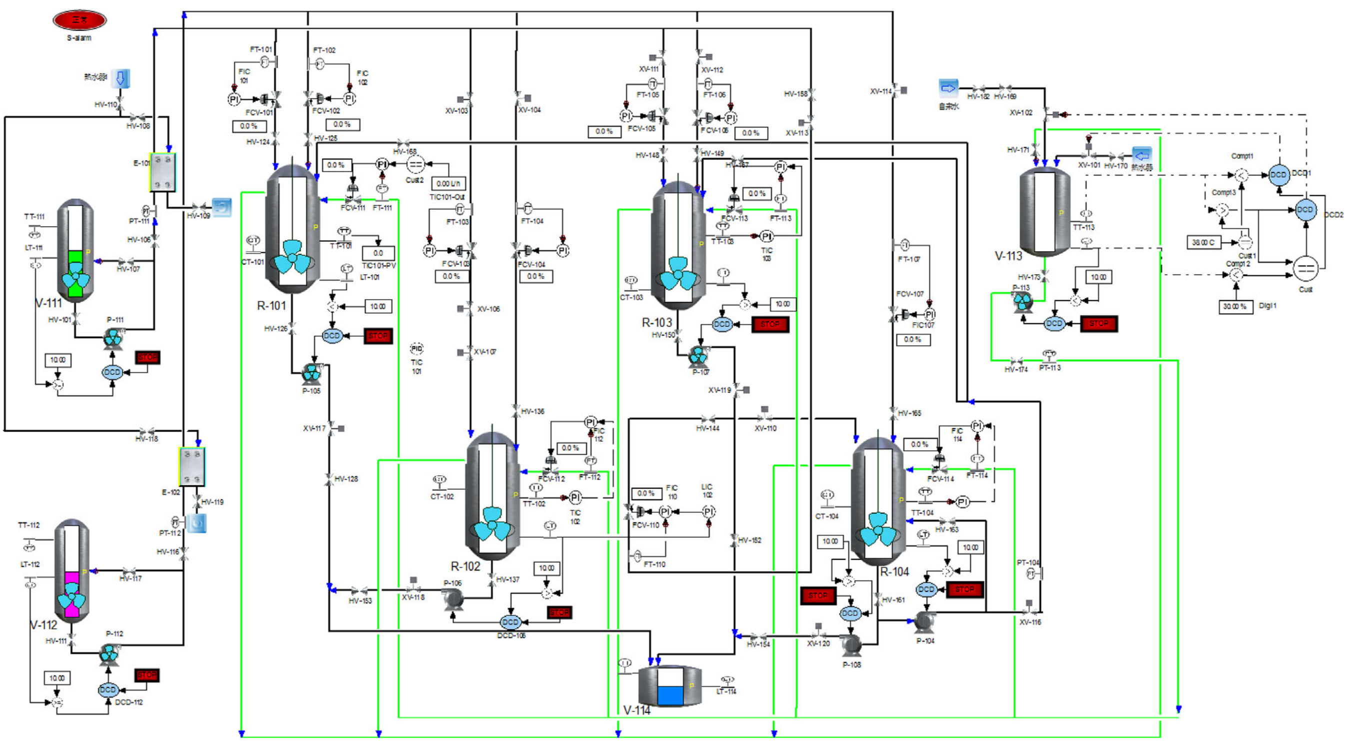

FIGURE 16 Experimental operation interface

of the step disturbance (31). The proposed approach converges to the desired trajectory more quickly. The ITSE for the two control methods is listed in Table 3. It is clear that the ITSE of MFASMC-L1 is generally smaller than that of MFASMC-L2, which indicates that the proposed method is robust.

5.2 | Experiment

The effectiveness of the proposed approach will now be validated using an experimental rig The Process Modelling and Control Group at the China University of Petroleum (East China) have installed an experimental rig which is shown in Figure 15. The operation interface of the rig is shown in Figure 16. The four reactors, labelled R101, R102, R103 and R104, can be connected in numerous ways for controller validation and testing (series, parallel, series and parallel). The process can implement continuous operation as well as enable measurement and control of the flow, liquid level and temperature. Only one reactor is used in this experiment. The sampling period is selected to be T s $= 1$ s and the sampling length is $N = 7000$. For reactor R101, the temperature is used as the system output

TABLE 4 Definitions for the variables of the process parameters

Variables	Definitions
F_0	Flow rate of fresh A
F_r	Flow rate of recycled A from reactor 2
F_3	Flow rate of additional fresh stream feeding pure A
F_1, F_2	Effluent flow rate from reactors 1, 2
C_{A1}, C_{A2}	Molar concentration of A in reactors 1, 2
T_1, T_2	Temperatures in reactors 1, 2
T_0, T_{03}	Feed stream temperatures to reactor 1, 2
Q_{r1}, Q_{r2}	Heat input rate into reactors 1, 2
C_{A0}, C_{A3}	Inlet reactant concentration of reactors 1, 2
V_1, V_2	Reactor volume of reactors 1, 2
$\Delta H_j, k_j, E_j$	Enthalpies, pre-exponential constants and activation energies of the reaction
ρ_s, R, c_p	Heat capacity, gas constant and density of fluid in the reactor

and the desired set-point is 30°C. The initial temperature of R101 is 26°C. The hot water flow in the jacket is used as the control input. The definitions of the variables and the values of the process parameters can be found in Tables 4 and 5, respectively. The MFASMC-L1 algorithm is used for control of the system.

TABLE 6 Controller parameters

MFASMC-L1	$\tau = 0.1, \chi = 0.1, \bar{\lambda} = 0.1, \bar{\gamma}_i = 0.1, q = 1, \varkappa = 1, T = 0.01, \alpha = 1.2$
MFASMC-L2	$\tau = 0.1, \chi = 0.1, \lambda = 0.1, \gamma_i = 0.1, q = 1, \varkappa = 1, T = 0.01, \rho = 0.04, a = 4.5, b_1 = 0.01, b_2 = 0.1$

For comparison purposes, the same experiment is performed using a classical MFASMC scheme (MFASMC-L2) [17]. Figure 17 shows the tracking performance of MFASMC-L2 and MFASMC-L1. It can be seen that the overshoot using MFASMC-L1 is smaller and the convergence speed is faster than obtained using MFASMC-L2. The experimental results further validate the proposed approach.

The parameters of the two control laws are listed in Table 6. The iteration accuracy is given by $\varepsilon^* = 0.001$ and $\varepsilon = 0.001$ and $\delta = 0.001$ using MFASMC-L1.

6 | CONCLUSIONS

This paper proposed a linearized Bregman iteration for model-free adaptive sliding mode controller design for unknown

TABLE 5 Values of the process parameters

$F_0 = 5.04$ m^3/h	$T_0 = 300$ K	$k_{30} = 3 \times 10^5$ h^{-1}	$C^s_{A1} = 1.67$ kmol/m^3
$F_1 = 50$ m^3/h	$T_{03} = 300$ K	$E_1 = 5 \times 10^4$ kJ/kmol	$T^s_2 = 431.91$ K
$F_3 = 30$ m^3/h	$\Delta H_1 = -5 \times 10^4$ kJ/kmol	$E_2 = 6.5 \times 10^4$ kJ/kmol	$C^s_{A2} = 1.73$ kmol/m^3
$F_r = 35$ m^3/h	$\Delta H_2 = -5.2 \times 10^4$ kJ/kmol	$E_3 = 5.5 \times 10^4$ kJ/kmol	$Q^s_1 = 1.2 \times 10^6$ kJ/h
$V_1 = 1.0$ m^3	$\Delta H_3 = -5.4 \times 10^4$ kJ/kmol	$c_p = 0.88$ kJ/kg K	$C^s_{A0} = 4$ kmol/m^3
$V_2 = 1.0$ m^3	$k_{10} = 3 \times 10^5$ h^{-1}	$\rho_s = 1000$ kg/m^3	$Q^s_2 = 1.6 \times 10^6$ kJ/h
$R = 8.314$ kJ/kmol K	$k_{20} = 3 \times 10^5$ h^{-1}	$T^s_1 = 469.2580$ K	$C^s_{A01} = 2$ kmol/m^3

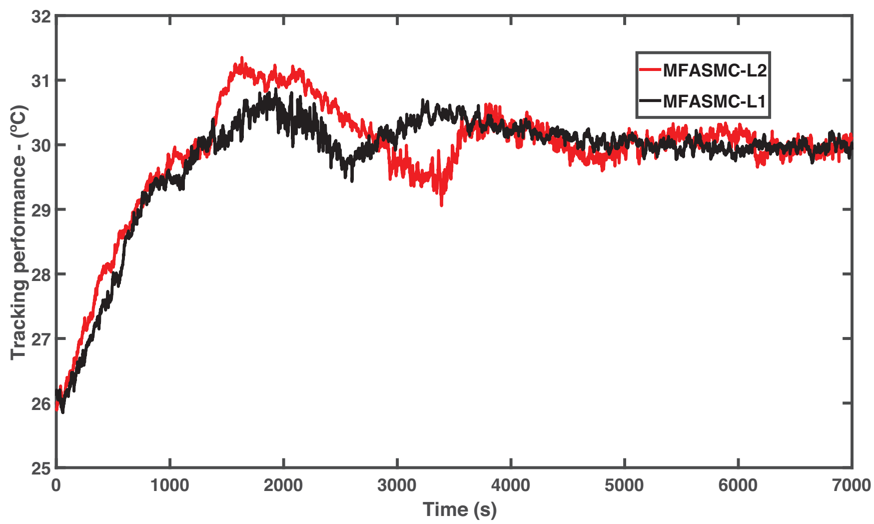

FIGURE 17 Tracking performance

MIMO non-linear discrete-time systems. Unlike traditional MFASMC, the proposed approach avoids matrix inversion and uses the L_1 norm in the controller design, which can enhance both computational efficiency and robustness. The proposed algorithm is straightforward to implement and improves system performance. Extensive simulation and experimental testing demonstrates the advantages of the proposed approach when compared to existing approaches in the literature. Future research efforts will aim to further improve the proposed solution with respect to robustness and consider further industrial applications.

ACKNOWLEDGEMENTS

This work is partially supported by the National Nature Science Foundation of China (61973315) and the Changjiang Scholars Program.

REFERENCES

1. Hou, Z., Wang, Z.: From model-based control to data-driven control: Survey, classification and perspective. Inform. Sci. 235, 3–35 (2013)
2. Silva, G.J., et al.: New results on the synthesis of PID controllers. IEEE Trans. Autom. Control 47(2), 241–252 (2002)
3. Meng, W., et al.: Robust iterative feedback tuning control of a compliant rehabilitation robot for repetitive ankle training. IEEE/ASME Trans. Mechatron. 22(1), 173–184 (2017)
4. Roman, R.C., et al.: Multi-input–multi-output system experimental validation of model-free control and virtual reference feedback tuning techniques. IET Control Theory Appl. 10(12), 1395–1403 (2016)
5. Şahin, S., Güzeliş, C.: Online learning ARMA controllers with guaranteed closed-loop stability. IEEE Trans. Neural Networks Learn. Syst. 27(11), 2314–2326 (2016)
6. He, W., et al.: Adaptive boundary iterative learning control for an Euler–Bernoulli beam system with input constraint. IEEE Trans. Neural Networks Learn. Syst. 29(5), 1539–1549 (2018)
7. Meng, D., Moore, K.L.: Convergence of iterative learning control for SISO nonrepetitive systems subject to iteration-dependent uncertainties. Automatica 79, 167–177 (2017)
8. Ahn, H.S., et al.: Iterative learning control: Brief survey and categorization. IEEE Trans. Syst. Man Cybern. Part C 37(6), 1099–1121 (2007)
9. Hou, Z., Jin, S.: Data-driven model-free adaptive control for a class of MIMO nonlinear discrete-time systems. IEEE Trans. Neural Networks 22(12), 2173–2188 (2011)
10. Hou, Z., Jin, S.: Model-Free Adaptive Control of MIMO Discrete-Time Nonlinear Systems.: Model Free Adaptive Control: Theory and Applications. CRC Press, Boca Raton, FL (2013)
11. Zhang, H., et al.: Optimal guaranteed cost sliding mode control for constrained-input nonlinear systems with matched and unmatched disturbances. IEEE Trans. Neural Networks Learn. Syst. 29(6), 2112–2126 (2018)
12. Fei, J., Lu, C.: Adaptive sliding mode control of dynamic systems using double loop recurrent neural network structure. IEEE Trans. Neural Networks Learn. Syst. 29(4), 1275–1286 (2018)
13. Fan, Q., Yang, G.: Adaptive actor-critic design-based integral sliding-mode control for partially unknown nonlinear systems with input disturbances. IEEE Trans. Neural Networks Learn. Syst. 27(1), 165–177 (2016)
14. Yue, M., et al.: MPC motion planning-based sliding mode control for underactuated WPS vehicle via Olfati transformation. IET Control Theory Appl. 12(4), 495–503 (2017)
15. Pang, B., Zhang, Q.: Sliding mode control for polynomial fuzzy singular systems with time delay. IET Control Theory Appl. 12(10), 1483–1490 (2018)
16. Wang, W., Hou, Z.: New adaptive quasi-sliding mode control for nonlinear discrete-time systems. J. Syst. Eng. Electron. 19(1), 154–160 (2008)
17. Wang, X., et al.: Data-driven model-free adaptive sliding mode control for the multi degree-of-freedom robotic exoskeleton. Inform. Sci. 327, 246–257 (2016)
18. Weng, Y., Gao, X.: Adaptive sliding mode decoupling control with data-driven sliding surface for unknown MIMO nonlinear discrete systems. Circuits Systems Signal Process. 36(3), 969–997 (2017)
19. Cai, J., et al.: Linearized Bregman iterations for compressed sensing. Math. Comput. 78(267), 1515–1536 (2009)
20. Figueiredo, M.A., et al.: Majorization–Minimization algorithms for wavelet-based image restoration. IEEE Trans. Image Process. 16(12), 2980–2991 (2007)
21. Wright, J., et al.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. 31(2), 210–227 (2009)
22. Sarpturk, S., et al.: On the stability of discrete-time sliding mode control systems. IEEE Trans. Autom. Control 32(10), 930–932 (1987)
23. Meng, D., et al.: Improve robustness of sparse PCA by L1-norm maximization. Pattern Recognit. 45(1), 487–497 (2012)
24. Horn, R.A., Johnson, C.R.: Norms for vectors and matrices. Matrix Analysis. Cambridge University Press, Cambridge (2005)
25. Zhang, J., et al.: Output feedback control of a class of discrete MIMO nonlinear systems with triangular form inputs. IEEE Trans. Neural Networks 16(6), 1491–1503 (2005)

How to cite this article: Gao S, Zhao D, Yan X, Spurgeon SK. Linearized Bregman iteration based model-free adaptive sliding mode control for a class of non-linear systems. *IET Control Theory Appl.* 2021;15:281–296. https://doi.org/10.1049/cth2.12039

APPENDIX A:

Proof of Lemma 4 is given as follows:
Proof. Consider Definition 1, the definition of Bregman distance implies that

$$H\left(\Delta u_j(k)\right) - H\left(\Delta u_{j-1}(k)\right) - \left\langle \Delta u_j(k) - \Delta u_{j-1}(k), U_j(k)\right\rangle \geq 0 \tag{A.1}$$

and

$$H\left(\Delta u_{j-1}(k)\right) - H\left(\Delta u_j(k)\right) - \left\langle \Delta u_{j-1}(k) - \Delta u_j(k), U_{j+1}(k)\right\rangle \geq 0, \tag{A.2}$$

where $H(\Delta u_j(k)) = \Delta u_j(k) - v(k)$.

By combining Equations (A.1) and (A.2)

$$\left\langle \Delta u_j(k) - \Delta u_{j-1}(k), U_{j+1}(k) - U_j(k)\right\rangle \geq 0. \tag{A.3}$$

Let $\theta \in (0, \frac{\alpha}{2\lambda})$ be a constant, define $\psi_j^\theta(k) = \langle \Delta u_j(k) - \Delta u_{j-1}(k), \lambda(U_{j+1}(k) - U_j(k)) + \theta(\Delta u_j(k) - \Delta u_{j-1}(k))\rangle$, then $\psi_j^\theta(k) \geq \theta\|\Delta u_j(k) - \Delta u_{j-1}(k)\|^2$. When $\Delta u_j(k) \neq \Delta u_{j-1}(k)$, $\psi_j^\theta(k) > 0$, and hence, a matrix can be defined as

$$Q_j^\theta(k) = \frac{1}{\psi_j^\theta}\left[\lambda\left(U_{j+1}(k) - U_j(k)\right) + \theta\left(\Delta u_j(k) - \Delta u_{j-1}(k)\right)\right] \times \left[\lambda\left(U_{j+1}(k) - U_j(k)\right) + \theta\left(\Delta u_j(k) - \Delta u_{j-1}(k)\right)\right]^T.$$

$$\tag{A.4}$$

It is obvious that $Q_j^\theta(k)$ is symmetric positive semi-definite and satisfies

$$\lambda\big(U_{j+1}(k) - U_j(k)\big) + \theta\big(\Delta u_j(k) - \Delta u_{j-1}(k)\big) = Q_j^\theta(k)\big(\Delta u_j(k) - \Delta u_{j-1}(k)\big). \tag{A.5}$$

Furthermore,

$$\begin{aligned}\left\|Q_j^\theta(k)\right\| &= \frac{1}{\psi_j^\theta(k)}\langle\lambda(U_{j+1}(k) - U_j(k)) + \theta(\Delta u_j(k) - \Delta u_{j-1}(k)),\\ &\quad\times\lambda(U_{j+1}(k) - U_j(k)) + \theta(\Delta u_j(k) - \Delta u_{j-1}(k))\rangle\\ &= \frac{1}{\psi_j^\theta(k)}(\langle\lambda(U_{j+1}(k) - U_j(k)) + \theta(\Delta u_j(k) - \Delta u_{j-1}(k)),\\ &\quad\times\lambda(U_{j+1}(k) - U_j(k))\rangle + \theta\psi_j^\theta(k))\\ &\le \frac{1}{\psi_j^\theta(k)}(\beta\lambda+\theta)\langle\Delta u_j(k) - \Delta u_{j-1}(k),\lambda\\ &\quad\times(U_{j+1}(k) - U_j(k))\rangle + \theta\\ &\le \frac{1}{\psi_j^\theta(k)}(\beta\lambda+\theta)\langle\Delta u_j(k) - \Delta u_{j-1}(k),\\ &\quad\times\lambda(U_{j+1}(k) - U_j(k)) + \theta(\Delta u_j(k) - \Delta u_{j-1}(k))\rangle + \theta\\ &= \beta\lambda + 2\theta.\end{aligned}$$

Note that the first equation in (14) yields

$$\lambda(U_{j+1}(k) - U_j(k)) + 2\alpha(\Delta u_j(k) - \Delta u_{j-1}(k)) = -\alpha(\Delta u_{j-1}(k) - v(k)). \tag{A.6}$$

By substituting (A.5) into (A.6)

$$(Q_j^\theta(k) + (2\alpha-\theta)I)(\Delta u_j(k) - \Delta u_{j-1}(k)) = -\alpha(\Delta u_{j-1}(k) - v(k)). \tag{A.7}$$

It follows that

$$\Delta u_j(k) - \Delta u_{j-1}(k) = -\frac{1}{2}\left(\frac{1}{2\alpha}Q_j^\theta(k) + \left(1 - \frac{1}{2\alpha}\theta\right)I\right)^{-1} \times(\Delta u_{j-1}(k) - v(k)). \tag{A.8}$$

Therefore,

$$\Delta u_j(k) - v(k) = \left(I - \frac{1}{2}\left(\frac{1}{2\alpha}Q_j^\theta(k) + \left(1 - \frac{1}{2\alpha}\theta\right)I\right)^{-1}\right) \times\big(\Delta u_{j-1}(k) - v(k)\big). \tag{A.9}$$

Since $0 \le Q_j^\theta(k) \le (\beta\lambda + 2\theta)I$, it can be established that

$$\left(1 - \frac{1}{2\alpha}\theta\right)I \le \frac{1}{2\alpha}Q_j^\theta(k) + \left(1 - \frac{1}{2\alpha}\theta\right)I \le \left(1 + \frac{1}{2\alpha}\beta\lambda + \frac{1}{2\alpha}\theta\right)I. \tag{A.10}$$

It follows that

$$\frac{\alpha}{2\alpha-\theta}I \ge \frac{1}{2}\left(\frac{1}{2\alpha}Q_j^\theta(k) + \frac{2\alpha-\theta}{2\alpha}I\right)^{-1} \ge \frac{\alpha}{2\alpha+\beta\lambda+\theta}I \tag{A.11}$$

so that

$$\left(1 - \frac{\alpha}{2\alpha-\theta}\right)I \le I - \frac{1}{2}\left(\frac{1}{2\alpha}Q_j^\theta(k) + \frac{2\alpha-\theta}{2\alpha}I\right)^{-1} \le \left(1 - \frac{\alpha}{2\alpha+\beta\lambda+\theta}\right)I \tag{A.12}$$

and

$$\left\|I - \frac{1}{2}\left(\frac{1}{2\alpha}Q_j^\theta(k) + \frac{2\alpha-\theta}{2\alpha}I\right)^{-1}\right\| \le \max \times\left\{\left\|1 - \frac{\alpha}{2\alpha-\theta}\right\|, \left\|1 - \frac{\alpha}{2\alpha+\beta\lambda+\theta}\right\|\right\}. \tag{A.13}$$

Considering Equation (A.9)

$$\|\Delta u_j(k) - v(k)\| \le \max\left\{\left\|1 - \frac{\alpha}{2\alpha-\theta}\right\|, \times\left\|1 - \frac{\alpha}{2\alpha+\beta\lambda+\theta}\right\|\right\}\|\Delta u_{j-1}(k) - v(k)a\| \tag{A.14}$$

Let $\eta = \max\{\|1 - \frac{\alpha}{2\alpha-\theta}\|, \|1 - \frac{\alpha}{2\alpha+\beta\lambda+\theta}\|\}$. Since θ can be any positive number in $(0, \frac{\alpha}{2\lambda})$, and the norm is a continuous function of θ, let $\theta \to 0$. From (A.10), it follows that

$$\eta \le \max\left\{\left\|1 - \frac{\alpha}{2\alpha}\right\|, \left\|1 - \frac{\alpha}{2\alpha+\beta\lambda}\right\|\right\}. \tag{A.15}$$

It follows that

$$0 < \frac{1}{2} < \|1 - \frac{\alpha}{2\alpha+\beta\lambda}\| < 1. \tag{A.16}$$

The above inequality also implies $\|\Delta u_j(k) - v(k)\| \le \eta\|\Delta u_{j-1}(k) - v(k)\|$, where $0 < \eta < 1$. □