Terminal Iterative Learning Scheme for Consensus Problem in Multi-Agent Systems with State Constraints

Yan Zhang\textsuperscript{1a}, Zijian Luo\textsuperscript{1b*}, Wenjun Xiong\textsuperscript{1c}

\textsuperscript{1}School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu 611130, China
\textsuperscript{a}karl_2002@163.com, \textsuperscript{b*zjluomath@126.com, c}xiongwenjun@swufe.edu.cn

Abstract—In this paper, we investigate the consensus problem of multi-agent systems with state constraints. To achieve the consensus effectively, the terminal iterative learning approach is proposed. This learning strategy is designed without the tracking error. And the consensus state is obtained by the information interaction between agents. Meanwhile, the constraint condition holds in terms of our learning strategy. It shows the consensus conditions ensure the achievement of the constraints. Finally, a numerical simulation is given to illustrate the effectiveness of the main results.

1. Introduction
Over the past decades, distributed coordination problems of multi-agent systems (MASs) have been studied because of their wide applications. MASs has several advantages, such as scalability, privacy protection and reliability [1-3]. As a kind of network systems, MASs attract much attention to the research on consensus, multi-robot systems, distributed optimization [4-6], etc. To analyse the consensus problems, the convergence rate plays a key role. However, these algorithms are complex, it brings some obstacles to theoretical analysis and practical applications. Hence, the simple and feasible control strategies should be designed to achieve the consensus.

As far as we know, iterative learning control (ILC) is an effective technology in the filed of industrial production. This control technique is widely applied in freeway traffic ramp metering [7], network systems [8] and so on. The tracking problem of MASs is also studied in terms of the ILC approach, such as the consensus tracking [9], optimal and adaptive learning [10,11], etc. The above results are obtained over the whole trial length. However, there is a class of applications that focus on the tracking at specific time instants [12-14]. Moreover, the terminal constraint is not considered in these researches. In some situations, the states of the system are usually constrained. Thus, it is meaningful to consider the terminal consensus problem with state constraints.

From the above discussion, the consensus problem is studied with state constraints by means of the terminal ILC strategies. The contributions of this paper are as follows: 1) Compared with traditional ILC algorithms, the consensus state can be obtained without the tracking error, it is achieved only by the information interaction between agents. 2) The results show that the consensus condition ensure the achievement of the constraints, and the state constraints hold at each iteration. Moreover, the sufficient conditions are obtained according to the inequalities skills.

The rest of the paper is organized as follows: Necessary notations and the problem formulation are proposed in Section 2. The main results are proposed in Section 3. A simulation result is presented in Section 4. Finally, the conclusion is shown in Section 5.
2. Preliminaries and Problem Formulation

2.1 Preliminaries

Notations: \( T_N = \{1, 2, \cdots, N\} \). \( \mathbf{0} \) and \( \mathbf{I} \) denote the zero vector and the unit matrix with appropriate dimension, respectively. \([0, T]\) represents time intervals and \( T \) is the fixed terminal time. Let \( \mathbf{G} = (\mathbf{V}, \mathbf{E}) \) be a weighted directed graph with the set of vertices \( \mathbf{V} = \{v_i : i \in T_N\} \), and the set of edges \( \mathbf{E} \subseteq \mathbf{V} \times \mathbf{V} \). A directed edge from \( v_i \) to \( v_j \) is denoted by an ordered pair \((i, j) \in \mathbf{E}\), which means agent \( v_j \) can receive the information from agent \( v_i \). The neighborhood of the agent \( v_i \) is denoted as \( \mathcal{N}_i = \{ v_j \in \mathbf{V} : (j, i) \in \mathbf{E}\} \). The weighted adjacency matrix is denoted as \( \mathbf{A} = (\alpha_{ij}) \in \mathbb{R}^{N \times N} \). And \( \alpha_{in} = 0 \) if \((j, i) \in \mathbf{E}\), and \( \alpha_{ij} = 0 \) otherwise. The Laplace matrix of \( \mathbf{G} \) is defined as \( \mathcal{L} = \{l_{ij} \} \in \mathbb{R}^{N \times N} \), where \( l_{ij} = -\alpha_{ij} \) if \( i \neq j \) and \( l_{ii} = \sum_{i \neq j} \alpha_{ij} \). A spanning tree is a directed graph, which has exactly one root vertex. Other vertexes are the child nodes of the root vertex. Then a graph has a spanning tree if \( \mathbf{V} \) and a subset of \( \mathbf{E} \) can form a tree.

2.2 Problem formulation

Consider the model which is studied in [13]. Let the interaction topology be described by graph \( \mathbf{G} \), and the dynamics of the agent \( i \) is described as

\[
\dot{x}_{k,i}(t) = u_{k,i}, \quad t \in [0, T], \quad i \in T_N, \tag{1}
\]

with initial state \( x_{k,i}(0) \). \( k \) denotes the kth iteration, \( \dot{x}_{k,i}(t) \) represents the derivative of \( x_{k,i}(t) \), \( x_{k,i}(t) \in \mathbb{R} \) is the state of agent \( i \) at time \( t \), and \( u_{k,i} \in \mathbb{R} \) is the control input of agent \( i \). In some situations, the state of the system are usually constrained. Thus, the constraints are as follows:

\[
b_z \leq x_{k,i}(t) \leq B_z, \quad t \in [0, T], \tag{2}
\]

where \( b_z, B_z \) are constant. Thus, the objective of the system (1) with constraints (2) is to achieve the consensus at the terminal time, i.e.

\[
\lim_{k \to \infty} x_{k,i}(T) = x_c, \quad \forall \ i \in T_N, \tag{3}
\]

where \( x_c \in \mathbb{R} \) is prescribed as a stable equilibrium at the terminal time \( T \).

The consensus objective requires the state of agents converge to the terminal time as the iteration number increases. Thus, it is necessary to design terminal iterative learning protocols. According to the principle of ILC and [13], the iterative learning strategy is designed as follows:

\[
\begin{align*}
\{ u_{0,i} \text{ is an appropriate initial input}, & \quad k = 0, \\
\{ u_{k+1,i} &= u_{k,i} + \gamma_i \sum_{j \neq i} \alpha_{ij} [x_{k,j}(T) - x_{k,i}(T)], & k \geq 1 \}, \tag{4}
\end{align*}
\]

where \( \alpha_{ij} \) is the \((i,j)\) entry of the adjacency matrix \( \mathbf{A} \), and \( \gamma_i \) is a positive learning gain parameter. If the information of the desired state \( x_d \in [b_z, B_z] \) can be obtained by agent \( v_i \), \( w_i > 0 \); otherwise \( w_i = 0 \). And the strategy (4) is redesigned as

\[
\begin{align*}
\{ u_{0,i} \text{ is an appropriate initial input}, & \quad k = 0, \\
\{ u_{k+1,i} &= u_{k,i} + \gamma_i \left\{ \sum_{j \neq i} \alpha_{ij} [x_{k,j}(T) - x_{k,i}(T)] + w_i [x_d - x_{k,i}(T)] \right\}, & k \geq 1 \}. \tag{5}
\end{align*}
\]

Due to lack of the empirical data, initial control input \( u_{0,i} \) can be appropriately selected to achieve the initial operation. Thus, the initial control does not satisfy the learning law. Moreover, the following assumption holds in this paper.

Assumption 1. The initial state satisfies \( x_{k,i}(0) \neq x_i(0) \in [b_z, B_z], \quad \forall \ i \in T_N \).
3. Main Results

3.1 Related lemmas

Lemma 1. ([13]) Consider the network (1) and (4) with a directed graph $G$, let the positive learning gain satisfy $T \gamma_i \sum_{j \in N_i} \alpha_{ij} < 1$, $i \in T_N$. Then the consensus objective (3) can be achieved asymptotically as $k \to \infty$ if and only if $G$ has a spanning tree.

Lemma 2. ([13]) Consider the network (1) and (5) with a directed graph $G$, let the positive learning gain satisfy $T \gamma_i \sum_{j \in N_i} \alpha_{ij} + w_i < 1$, $i \in T_N$. Then the consensus objective (3) is achieved asymptotically as $k \to \infty$ and $x_e = z_d$ if and only if the matrix $L + W$ is non-singular, where $W = \text{diag}\{w_1, w_2, \ldots, w_N\}$.

3.2 Consensus analysis

Theorem 1. With Assumption 1, assume the systems (1) and (4) with a digraph $G$ has a spanning tree, and the initial input satisfies $b_x - x_i(0) \leq T b_{0,i} \leq B_x - x_i(0)$. One can choose the positive learning gain to satisfy the inequality of Lemma 1, then condition (2) holds at each iteration, and the consensus objective (3) can be achieved asymptotically as $k \to \infty$.

Proof. From (1), it is easy to know that $x_{k+1}(t) = x_k(t) + tu_k(t)$. According to Assumption 1 and learning strategy (4), the state of agent $i$ at the terminal time $T$ satisfies

$$x_{k+1}(T) = x_k(T) + T(u_{k+1,i} - u_{k,i})$$

$$= x_k(T) + \sum_{j \in N_i} \alpha_{ij} [x_j(T) - x_k(T)]$$

$$= \left(1 - \gamma_i \sum_{j \in N_i} \alpha_{ij}\right)x_k(T) + T \gamma_i \alpha_{ij} x_{k+1}(T) + T \gamma_i \alpha_{ij} x_{k+2}(T) + \cdots + T \gamma_i \alpha_{ij} x_{k+N}(T).$$

Let $\lambda_i = 1 - \gamma_i \sum_{j \in N_i} \alpha_{ij}$, $\lambda_2 = T \gamma_i \alpha_{ij}$, $\cdots$, $\lambda_{N+1} = T \gamma_i \alpha_{ij}$, it is easy to know $\sum_{i=1}^{N+1} \lambda_i = 1$. Since $T \gamma_i \sum_{j \in N_i} \alpha_{ij} < 1$, $i \in T_N$, $\gamma_i > 0$ and $\alpha_{ij} \geq 0$, one can get $0 < \lambda_i < 1$. Thus, $x_{k+1}(t)$ is the convex combination of states of its neighbor nodes at $k$th iteration.

Since $b_x - x_i(0) \leq T b_{0,i} \leq B_x - x_i(0)$, the appropriate initial state $x_i(0)$ and control input $u_{0,i}$ is chosen to satisfy $x_{0,i}(T) \in [b_x, B_x]$, $\forall \, i \in T_N$. According to (6), $x_i(T)$ is the convex combination of $x_{0,j}(T)$, $j \in N_i$. Note the $[b_x, B_x]$ is a convex set, and a set $S$ is convex if and only if any convex combination of points in $S$ belongs to $S$. One can know that $x_i(T) \in [b_x, B_x]$, $\forall \, i \in T_N$. According to mathematics inductive method, it is not difficult to know that $x_{k,i}(T) \in [b_x, B_x]$, $\forall \, i \in T_N$, $k \in \mathbb{N}^+$. Further, $x_{k,i}(T) = x_{k,i}(0) + T u_{k,i}$. Hence, one can obtain

$u_{k,i} = \left[ b_x - x_i(0) / T, B_x - x_i(0) / T \right]$, $\forall \, i \in T_N$, $k \in \mathbb{N}^+$.

Note $x_{k,i}(t) = x_{k,i}(0) + t u_{k,i}$, then

$$b_x \leq \frac{t}{T} (b_x - x_i(0)) + x_i(0) \leq x_i(0) + t u_{k,i} \leq \frac{t}{T} (B_x - x_i(0)) + x_i(0) \leq B_x.$$  

(7)

Inequality (7) means $b_x \leq x_{k,i}(t) \leq B_x$, i.e. the condition (2) holds at each iteration. The consensus objective (3) can be immediately achieved from Lemma 1.
Remark 1. Compared with typical ILC Schemes [8-10], the desired state information is not used in our terminal learning strategy. And the consensus state can be obtained by the information interaction between agents.

Theorem 2. With Assumption 1, assume the systems (1) and (5) with a digraph $G$ has a spanning tree, the matrix $L + \mathbf{W}$ is non-singular and the initial input satisfies $b_x - x_i(0) \leq T u_{0,i} \leq B_x - x_i(0)$.

One can choose positive learning gain to satisfy the inequality of Lemma 2, then condition (2) holds at each iteration, the consensus objective (3) is achieved asymptotically as $k \to \infty$ and $x_i = x_d$.

Proof. The proof is similar to Theorem 1. Set $\epsilon_{k,i}(T) = x_{k,i}(T) - x_d$, linking with (6), one has

$$
\epsilon_{k+1,i}(T) = \epsilon_{k,i}(T) + T (u_{k+1,i} - u_{k,i})
$$

$$
= \epsilon_{k,i}(T) + T \gamma_i \left\{ \sum_{j \in \mathcal{N}_i} a_{ij} [\epsilon_{k,j}(T) - \epsilon_{k,i}(T)] - w_i \epsilon_{k,i}(T) \right\}
$$

$$
= \left( 1 - T \gamma_i \sum_{j \in \mathcal{N}_i} a_{ij} + w_i \right) \epsilon_{k,i}(T) + T \gamma_i \sum_{j \in \mathcal{N}_i} a_{ij} \epsilon_{k,j}(T). \tag{8}
$$

Similar to the proof of Theorem 1, it is not hard to know $x_{k,i}(T) \in [b_x, B_x], \forall \ i \in \mathcal{T}_N, \ k \in \mathcal{N}^+$. And the $x_{k,i}(t) \in [b_x, B_x]$ in line with (7), i.e. the condition (2) holds at each iteration. With Lemma 2, it is immediately to get $x_i = x_d$.

Remark 2. From Theorems 1 and 2, one knows that the consensus conditions of Lemmas 1 and 2 are sufficient to satisfy the constraints. Then the consensus problem with state constraints is effectively solved by the terminal ILC approach. And our learning strategy is efficient and simple.

4. Simulation Results

Considering the multi-agent systems (1), the communication graph $G$ of agents is shown in Figure 1(a), and the agents are denoted by nodes 1-6. Red lines represent the communication between the agents, and the blue numbers represent the weight of connected edges.

Set the initial states of agents 1-6 are -0.508, -0.752, -0.422, -0.838, -0.580 and -0.340, respectively. Let the state constraint bound $b_x = -0.9$ and $B_x = -0.3$. Let $t \in [0, T], T = 1$, i.e. the running time of each iteration is 1 second. Moreover, learning strategy (4) is adopted, set initial input $u_{0,i} = 0, \forall i \in \mathcal{T}_N$, the learning gain parameters $\gamma_1 = 2/3, \gamma_2 = 1/2, \gamma_3 = 1/3, \gamma_4 = 5/12, \gamma_5 = 2/5$ and $\gamma_6 = 1/4$. By calculation, it is easy to check that the conditions of Theorem 1 hold. Then constraint (2) holds at each iteration, and the consensus state can be achieved at the terminal time.

The result of terminal iterative learning is shown in Fig. 1(b)-(d). Fig. 1(b) shows the state at the 10th iteration, Fig.1(c) shows the state at the 50th iteration. From Fig. 1(b) and (c), one knows the constraint (2) holds, and the terminal consensus is asymptotically achieved. Fig. 1(d) shows terminal state of 1-100th iterations. It implies that the terminal state errors are decreased with the increasing iteration step size $k$. Thus, Fig.1 shows the consensus problem (3) is effectively achieved according to our learning strategy. And the validity of Theorem 1 is illustrated.
Theorem 2 has the similar conclusion, which is not repeated here. In simulation, the information of desired state is not used in our learning strategy. The consensus state is obtained by the interaction of agents, and the constraint (2) holds at the same time. It shows the effectiveness of the main results.

5. Conclusion
Based on the results and discussions presented above, the conclusions are obtained as below:

1) The consensus problem has been studied with the state constraints. And the state constraints have hold at each iteration. The main results have shown that the consensus condition is sufficient to satisfy the constraints.

2) The tracking error has not been used in our learning strategy (4), the terminal consensus state has been achieved by the the information interaction of agents. It implies our learning strategy is effective to solve the terminal consensus problem.

Acknowledgments
This work was jointly supported by the National Natural Science Foundation of China (61873344), the Fundamental Research Funds for the Central Universities (JBK190502).

References
[1] Manjula, K., Karthikeyan, P. (2010) Distributed computing approaches for scalability and high performance. Int. J. Eng. Sci. Technol., 2(6): 2328-2336.
[2] Léauté, T., Faltings, B. (2013) Protecting privacy through distributed computation in multi-agent decision making. J. Artif. Intell. Res., 47: 649-695.
[3] Ke, W., Wang, S. (1997) Reliability evaluation for distributed computing networks with imperfect nodes. IEEE Trans. Reliab., 46(3): 342-349.
[4] Radenković, M. S., Krstić, M. (2018) Distributed adaptive consensus and synchronization in complex networks of dynamical systems. Automatica, 91: 233–243.

[5] Chen, D., Liu, X., Yu, W. (2020) Finite-time fuzzy adaptive consensus for heterogeneous nonlinear multi-agent systems. IEEE Trans. Net. Sci. Eng., 7(4): 3057–3066.

[6] Zhao, Z., Chen, G., Dai, M. (2018) Distributed event-triggered scheme for a convex optimization problem in multi-agent systems. Neurocomputing, 284: 90-98.

[7] Chi, R., Hou, Z., Jin, S., Wang, D., Hao, J. (2013) A data-driven iterative feedback tuning approach of ALINEA for freeway traffic ramp metering with PARAMICS simulations. IEEE Trans. Ind. Inform., 9(4): 2310–2317.

[8] Xiong, W., Ho, D. W., Wen, S. (2021) A periodic iterative learning scheme for finite-iteration tracking of discrete networks based on FlexRay communication protocol. Inform. Sci., 548(16): 344–356.

[9] Luo, Z., Xiong, W., He, W., Chen, Y. (2020) Observer-based state tracking for discrete linear multi-agent systems with switching topologies via learning control strategies. IET Control Theory Appl., 14(12): 1639–1647.

[10] Yang, S., Xu, J.-X., Huang, D., Tan, Y. (2014) Optimal iterative learning control design for multi-agent systems consensus tracking. Syst. Contr. Lett., 64: 80–89.

[11] He, S., Fang, H., Zhang, M., Liu, F., Ding, Z. (2020) Adaptive optimal control for a class of nonlinear systems: the online policy iteration approach. IEEE Trans. Neur. Net. Lear. Syst., 31(2): 549–558.

[12] Han, J., Shen, D., Chien, C.-J. (2018) Terminal iterative learning control for discrete-time nonlinear systems based on neural networks. J. Franklin Ins., 35(8): 3641–3658.

[13] Meng, D., Jia, Y., Du, J. (2014) Finite-time consensus protocols for networks of dynamic agents by terminal iterative learning. Inter. J. Syst. Sci., 45(11): 2435–2446.

[14] Chen, B., Chu, B. (2019) Distributed norm optimal iterative learning control for point-to-point consensus tracking. IFAC Papers On Line, 52(29): 292–297. https://doi.org/10.1016/j.ifacol.2019.12.665