Cooperative Online Learning for Multi-Agent System Control via Gaussian Processes with Event-Triggered Mechanism

Xiaobing Dai[1], Zewen Yang[1*], *Member, IEEE*, Sihua Zhang[1,2], Di-Hua Zhai[2], Yuanqing Xia[2], *Fellow, IEEE*, Sandra Hirche[1], *Fellow, IEEE*

***Abstract*—In the realm of the cooperative control of multi-agent systems (MASs) with unknown dynamics, Gaussian process (GP) regression is widely used to infer the uncertainties due to its modeling flexibility of nonlinear functions and the existence of a theoretical prediction error bound. Online learning, which involves incorporating newly acquired training data into Gaussian process models, promises to improve control performance by enhancing predictions during the operation. Therefore, this paper investigates the online cooperative learning algorithm for MAS control. Moreover, an event-triggered data selection mechanism, inspired by the analysis of a centralized event-trigger, is introduced to reduce the model update frequency and enhance the data efficiency. With the proposed learning-based control, the practical convergence of the MAS is validated with guaranteed tracking performance via the Lyapunov theory. Furthermore, the exclusion of the Zeno behavior for individual agents is shown. Finally, the effectiveness of the proposed event-triggered online learning method is demonstrated in simulations.**

***Index Terms*—Learning-based control, cooperative learning, event-triggered learning, Gaussian processes, multi-agent system.**

I. Introduction

Cooperative control for the multi-agent system (MAS) with unknown models or environmental uncertainties has drawn large attention over the past two decades, particularly in fields such as aerial drones [1], underwater vehicles [2] and networked sensors [3]. To compensate for the uncertainties, machine learning methods are employed on each agent to learn the unknown components from collected data, and then subsequently integrate into the model-based controller design [4], [5]. Specifically, among machine learning techniques, Gaussian process (GP) regression [6] is popular for model estimation in safety-critical control [7]–[9] due to its capability of inferring unknown dynamics with a probabilistically guaranteed prediction performance [10].

*Corresponding author.

[1]Xiaobing Dai, Zewen Yang, Sihua Zhang and Sandra Hirche are with the Chair of Information-oriented Control (ITR), School of Computation, Information and Technology (CIT), Technical University of Munich (TUM), 80333 Munich, Germany (email: xiaobing.dai, zewen.yang, sihua.zhang, hirche@tum.de). [2]Sihua Zhang, Di-Hua Zhai and Yuanqing Xia are with the School of Automation, Beijing Institute of Technology, Beijing, People's Republic of China (email: sihua.zhang, zhaidih, xia_yuanqing@bit.edu.cn).

This work is supported by the Federal Ministry of Education and Research of Germany in the programme of "Souverän. Digital. Vernetzt." under the joint project 6G-life with identification number: 16KISK002, by the European Research Council (ERC) Consolidator Grant "Safe data-driven control for human-centric systems (CO-MAN)" under grant agreement number 864686, and National Natural Science Foundation of China under Grant 62173035.

The efficacy of learning-based control systems is usually contingent upon the accuracy of the predictions from GP models, which can be enhanced by leveraging more training data [11]. Therefore, cooperative learning is utilized to augment the inference accuracy while managing large datasets within MAS by dividing the overall data set into several sets of individual agents, where the predictions are aggregated via the communication networks. For instance, incorporating the communication graph among agents, a topology-aware aggregation method is first introduced in [12], inspired the product of experts (PoE) [13]. Moreover, several variants [14], [15] are introduced to reduce the computational complexity from cooperative learning. Additionally, applying consensus theory, a dynamics average consensus algorithm is incorporated into GP aggregation [16] to synchronize the individual predictions from each agent. However, the performance of cooperative learning with aggregation methods remains constrained by the precision of the local GP regression models.

To surmount this limitation and further improve prediction accuracy, the integration of online learning emerges as a promising strategy during system operation. Among online learning methods for MAS, collective online learning of GP is proposed to bolster the precision of the local GP model in each agent, by real-time optimization of hyper-parameters through the streaming data [17]. However, considering the design of this algorithm is strongly based on empirical approximations, the extension of the prediction error bound from the exact GP regression is blocked and thus its practical application in safety-critical scenarios is constrained. In addition to the above methods, online data collection yields more accurate prediction models by generating larger data sets [18], while maintaining the existence of the prediction error bound [19]. The online data collection for multi-agent system is introduced in [20] to achieve formation control, and in [21] for rigidity-based flocking control. Note that the prediction in [21], [22] uses merely local GP models on each agent, i.e., without cooperative learning. The combination of online and cooperative learning for control performance improvement of MAS, to the best knowledge of the authors, has not been addressed yet. Moreover, these works, while intuitive and practical for improving the local GP models, raise concerns about data storage and computational resource requirements due to the accumulated data. Furthermore, while the adopted time-trigger serves as an intuitive and practical strategy to reduce computational demands, it overlook the varying significance of

data with respect to control performance, often resulting in the collection of unnecessary data and diminished data efficiency.

In response to the concern of data efficiency, a smart data selection strategy becomes imperative, ensuring the exclusive collection of necessary data. Event-triggered data selection methods, recognized for their efficiency in control scenarios, offer benefits in terms of data storage and computational resources while maintaining desired control performance [23], [24]. The concept of event-triggered learning has been extensively explored in the control of MASs based on neural networks (NNs) [25]. The trigger mechanisms are designed to determine the instances for updating control inputs [26] or broadcasting the weights in the fully connected layer of NN for cooperative learning [27], aiming to alleviate communication burdens. While these methods are shown effective for NN-based control, they strongly depend on the unique regression form of NNs, i.e., linear regression with nonlinear features, which hampers a straightforward extension to other learning-based control in MAS, including GP-based methods [20]. In the realm of event-triggered online learning with GPs via data collection, studies have been conducted on single-system control using feedback linearization [28] and back-stepping [29]. However, these studies presume complete knowledge of system states, rendering them impractical for deployment in networked MASs. Therefore, for the robustness and scalability of the proposed event-trigger mechanism, the capability of distributed computation becomes pivotal. Although event-triggered learning mechanisms have been investigated in model-based and NN-based control, as far as we are aware, the study of effective GP-based online learning control using event-triggered data collection especially in distributed ways has not been explored in the existing literature.

A. Contribution and Structure

In this work, we develop online cooperative learning strategies for GP-based MAS control with event-triggered mechanisms. To begin, a learning-based leader-follower time-varying formation control framework for high-order MASs in directed topology with unknown dynamics is proposed, where the derived methodology extends naturally to other control tasks, such as consensus control. To improve the learning performance during the operation, a general cooperative online learning strategy based on aggregation and online data collection is proposed, and its prediction performance is analyzed. Furthermore, to enhance streaming data collection efficiency and alleviate computational burdens from prediction model updates, a distributed event-triggered online learning strategy is designed, which is inspired by the analysis of centralized approach. To obviate the requirement for a central node with access to all agents, we further propose a fully distributed event-triggered approach, which not only exhibits enhanced scalability but also entails fewer model updates, all while maintaining guaranteed prediction performance. Moreover, the achievement of desired control performance, i.e., ensuring an overall tracking error bound around the equilibrium point, is substantiated by using the proposed event-triggered online learning mechanisms. Additionally, a rigorous analysis is provided to show the exclusion of Zeno behavior for each agent within the MAS. Finally, the effectiveness of the proposed event-triggered online learning algorithm is demonstrated through simulations, which shows the enhancement of the control performance with less frequent model updates compared to time-triggered learning and offline learning.

The remainder of the paper is structured as follows: Section II outlines the problem setting. In Section III, cooperative online learning with Gaussian process regression is discussed for MAS control with their performance analysis. The centralized and distributed event-triggered online learning mechanisms are proposed in Section IV with the discussion of the Zeno behavior. Numerical simulations are presented in Section V to demonstrate the effectiveness of the proposed method. Finally, Section VI concludes the paper.

B. Notation and Graph Theory

The natural numbers with/without zero are denoted by $\mathbb{N}/\mathbb{N}_+$, real positive numbers with and without zero by $\mathbb{R}_{0,+}$ and $\mathbb{R}_+$, respectively. Minimum/maximum eigenvalues of a square matrix $\boldsymbol{A}$ are denoted by $\underline{\lambda}(\boldsymbol{A})/\bar{\lambda}(\boldsymbol{A})$. Unless explicitly specified, $|\cdot|$ refers to the element-wise absolute operator, and $\|\cdot\|$ represents the Euclidean norm. The i-th entry of a vector $\boldsymbol{a}$ is represented as a_i, whereas a_{ij} signifies the element at the intersection of the i-th row and the j-th column within matrix $\boldsymbol{A}$. The symbol $\otimes$ denotes the Kronecker product. The diagonal operator for scalar inputs denotes $\mathrm{diag}(\cdot)$, and the block diagonal operator for vector/matrix inputs denotes $\mathrm{blkdiag}(\cdot)$. The identity matrix with the dimension of $m \times m$ is denoted by $\boldsymbol{I}_m$, and the column vector $m \times 1$ vector with all components equal to one is denoted as $\mathbf{1}_m$.

The communication network for information exchange among $N \in \mathbb{N}_+$ agents is defined by a directed graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{1, \cdots, N\}$ is the vertex set representing the indices of agents and $\mathcal{E} \in \mathcal{V} \times \mathcal{V}$ is the edge set. The agent i can receive the information from agent j when $(j,i) \in \mathcal{E}, \forall i,j \in \mathcal{V}$. The topology of the graph is characterized by a weighted adjacency matrix denoted as $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{N\times N}$. The entries $a_{ij} > 0$ when there exists a communication channel from agent j to agent i, i.e., $(j,i) \in \mathcal{E}$, otherwise $a_{ij} = 0, \forall i,j \in \mathcal{V}$. Furthermore, the out-degree Laplacian matrix is defined as $\boldsymbol{\mathcal{L}} = \{l_{ij}\}_{i,j\in\mathcal{V}} \in \mathbb{R}^{N\times N}$ with $l_{ii} = \sum_{j=1}^{N} a_{ij}$ and $l_{ij} = -a_{ij}$ for $j \neq i$. The set $\mathcal{N}_i$ contains all the neighbor agents of agent i, i.e., $\mathcal{N}_i = \{j \in \mathcal{V} | a_{ij} > 0\}$, and the set $\bar{\mathcal{N}}_i$ is defined as $\bar{\mathcal{N}}_i = \{j \in \mathcal{V} | j \in \mathcal{N}_i \ \wedge \ i \in \mathcal{N}_j\}$ such that agent j can get $\boldsymbol{x}_i$ from agent i and vice versa.

II. Problem Setting and Preliminaries

A. Multi-Agent System

In this paper, we consider a MAS with $N \in \mathbb{N}_+$ homogeneous agents. The i-th agent follows a n-order continuous dynamical system with $n \in \mathbb{N}_+$ described as

$$\begin{aligned} \dot{\boldsymbol{x}}_{i,k} &= \boldsymbol{x}_{i,k+1}, \ \forall k = 1, \cdots, n-1, \\ \dot{\boldsymbol{x}}_{i,n} &= \boldsymbol{h}(\boldsymbol{x}_i) + \boldsymbol{g}(\boldsymbol{x}_i)\boldsymbol{u}_i + \boldsymbol{f}(\boldsymbol{x}_i), \ i \in \mathcal{V}, \end{aligned} \tag{1}$$

where the $p \in \mathbb{N}_+$ dimension states denote $\boldsymbol{x}_{i,j} \in \mathbb{R}^p, \forall j = 1, \cdots, n$. The concatenated system states and control input for

agent i denote $\boldsymbol{x}_i = [\boldsymbol{x}_{i,1}^T, \cdots, \boldsymbol{x}_{i,n}^T]^T \in \mathbb{X} \subset \mathbb{R}^{np}$ and $\boldsymbol{u}_i \in \mathbb{R}^q$ with $q \in \mathbb{N}_+$, respectively. The continuous function $\boldsymbol{h}(\cdot) : \mathbb{X} \to \mathbb{R}^n$ and the non-singular continuous function $\boldsymbol{g}(\cdot) : \mathbb{X} \to \mathbb{R}^{np\times q}$ represent the known parts of the system dynamics, which are usually obtained by using the first principle. Note that the non-singular $\boldsymbol{g}(\cdot)$ is a prerequisite for feedback controller design [30], which ensures each agent is controllable at any state in $\mathbb{X}$ [16]. The unmodeled part and external environmental uncertainties are encoded into function $\boldsymbol{f}(\cdot) : \mathbb{X} \to \mathbb{R}^p$, so it is considered as unknown but identical for each agent.

Remark 1: The homogeneous multi-agent system with identical unknown functions $\boldsymbol{f}(\cdot)$ considered in this paper allows cooperative learning, i.e., the employment of data-driven models from neighbors to enhance local prediction. For heterogeneous MAS with different unknown $\boldsymbol{f}_i(\cdot)$, cooperative learning by aggregating predictions from neighbors becomes meaningless. Such individual learning is covered in the proposed cooperative learning framework as a special case with no trust to neighbors' predictions. Considering the focus of this paper lies on the design of event-triggered cooperative online learning, the extension to heterogeneous MAS with different distributed control laws is left for future work.

In this paper, the dimensions of the state in every order, i.e., $\boldsymbol{x}_{i,k}$, and control input $\boldsymbol{u}_i$ are set as scalar, indicating $p=q=1$ for notational simplicity. The derived results can be directly extended to high-dimensional systems by using Kronecker product and multi-output machine learning methods.

The control objective is to achieve a leader-follower time-varying formation, where the dynamics of the leader follows

$$\dot{x}_{l,k} = x_{l,k+1},\ \forall k = 1, \cdots, n-1, \quad \dot{x}_{l,n} = x_{l,r}(t), \quad (2)$$

where the leader states denote $\boldsymbol{x}_l = [x_{l,1}, \cdots, x_{l,n}]^T \in \mathbb{R}^n$ with $x_{l,k} \in \mathbb{R}, \forall k = 1, \cdots, n$ and the known continuous function $x_{l,r}(\cdot) : \mathbb{R}_{0,+} \to \mathbb{R}$. Moreover, each agent tracks the leader but keeps a predefined relative time-varying distance as

$$\dot{s}_{i,k} = s_{i,k+1},\ \forall k = 1, \cdots, n-1, \quad \dot{s}_{i,n} = s_{i,r}(t) \quad (3)$$

for all $i \in \mathcal{V}$, where its concatenated states $\boldsymbol{s}_i = [s_{i,1}, \cdots, s_{i,n}]^T \in \mathbb{R}^n$ with $s_{i,k} \in \mathbb{R}, \forall k = 1, \cdots, n$ and the known continuous function $s_{i,r}(\cdot) : \mathbb{R}_{0,+} \to \mathbb{R}$. The structures of the leader and relative trajectory in (2) and (3) guarantee that all agents $i \in \mathcal{V}$ are capable of following its own reference $\boldsymbol{x}_l + \boldsymbol{s}_i$ [16]. Furthermore, the desired reference $\boldsymbol{x}_l + \boldsymbol{s}_i$ for each agent i and the derivative of the highest order of leader dynamics $x_{l,r}(\cdot)$ satisfy the following assumption.

Assumption 1: There exist well-defined positive constants $F_l, F_{r,i} \in \mathbb{R}_{0,+}$, such that $x_{l,r}(\cdot)$ and the derivative of the references $\dot{\boldsymbol{x}}_l + \dot{\boldsymbol{s}}_i$ are bounded as $|x_{l,r}(t)| \leq F_l$ and $\|\dot{\boldsymbol{x}}_l(t) + \dot{\boldsymbol{s}}_i(t)\| \leq F_{r,i}$ respectively for $\forall t \in \mathbb{R}_{0,+}$ and $\forall i \in \mathcal{V}$.

In practice, the dynamics of the leader $\boldsymbol{x}_l$ in (2) and relative distances $\boldsymbol{s}_i$ in (3) for $\forall i \in \mathcal{V}$ are designed such that $\boldsymbol{x}_l + \boldsymbol{s}_i \in \mathbb{X}$, indicating each entry in $\boldsymbol{x}_l + \boldsymbol{s}_i$ is bounded. Then, the boundness of its derivative, i.e., $\|\dot{\boldsymbol{x}}_l(t) + \dot{\boldsymbol{s}}_i(t)\|$, only requires bounded $x_{l,r}(\cdot)$ and $s_{i,r}(\cdot)$ for $\forall t \in \mathbb{R}_{0,+}$ considering the structures in (2) and (3). Since the dynamics of the leader and relative reference are design choices, Assumption 1 is not restrictive. It can be easily satisfied by choosing $x_{l,1}(t)$ and $s_{i,1}(t)$ as at least $(n+1)$-th smooth, i.e., $x_{l,1}(\cdot), s_{i,1}(\cdot) \in \mathcal{C}^{n+1}$.

Note that, while the the states $\boldsymbol{s}_i$ and $s_{i,r}(\cdot)$ are available for each agent i, only the agents connected to the leader are able to obtain the information of leader states $\boldsymbol{x}_l$ and $x_{l,r}(\cdot)$. In order to describe the connectivity between the leader and followers, a diagonal matrix $\boldsymbol{\mathcal{B}} = \mathrm{diag}(b_{11}, \cdots, b_{NN}) \in \mathbb{R}^{N\times N}$ is adopted, where $b_{ii} = 1$ indicates the agent i receives the information from the leader and $b_{ii} = 0$ otherwise. Due to the existence of leader disconnection, the controlled system cannot achieve asymptotically stability with formation error $\boldsymbol{\vartheta} = \boldsymbol{0}$, where

$$\boldsymbol{\vartheta} = \boldsymbol{x} - \boldsymbol{s} - \boldsymbol{1}_N \otimes \boldsymbol{x}_l \quad (4)$$

with $\boldsymbol{x} = [\boldsymbol{x}_1^T, \cdots, \boldsymbol{x}_N^T]^T$ and $\boldsymbol{s} = [\boldsymbol{s}_1^T, \cdots, \boldsymbol{s}_N^T]^T$. Instead, the formation error $\boldsymbol{\vartheta}$ is bounded by a positive constant $\bar{\vartheta} \in \mathbb{R}_{0,+}$, i.e., $\|\boldsymbol{\vartheta}\| \leq \bar{\vartheta}$. To achieve the cooperative control for the leader-follower formation task, the following assumption on the communication topology is required.

Assumption 2 ([31]): The augmented graph characterized by $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{B}}$ contains a spanning tree. Moreover, the leader is the root without incoming edge from the agents.

Assumption 2 is commonly found in MAS control to ensure the information of the leader is directly available by some agents, and propagates through the network [12], [16]. Moreover, Assumption 2 is essential to guarantee the achievement of leader-follower task and leads to the following lemma.

Lemma 1 ([32]): Suppose Assumption 2 holds and choose $\boldsymbol{q} \in \mathbb{R}^N$ and $\boldsymbol{P} \in \mathbb{R}^{N\times N}$ defined as

$$\boldsymbol{q} = [q_1, \cdots, q_N]^T = (\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}})^{-1}\boldsymbol{1}_N, \quad \boldsymbol{P} = \mathrm{diag}(q_1^{-1}, \cdots, q_N^{-1}).$$

Let $\boldsymbol{Q} = \boldsymbol{P}(\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}}) + (\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}})^T\boldsymbol{P} \in \mathbb{R}^{N\times N}$, then the matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ are symmetric positive definite.

Lemma 1 provides a choice of the Lyapunov function for high-order MAS [33] used for stability analysis in Section III.

B. Distributed Control Law

To achieve the leader-follower time-varying formation control task within the multi-agent framework, a distributed control structure for the individual agent i is proposed as

$$u_i = g^{-1}(\boldsymbol{x}_i)\big(\phi(\mathcal{J}_i, \mathcal{I}_i, \mathcal{J}_{r,i}) - h(\boldsymbol{x}_i) - \hat{f}_i(\boldsymbol{x}_i)\big),\ \forall i \in \mathcal{V}, \quad (5)$$

where the linear consensus law $\phi(\cdot, \cdot, \cdot)$ is designed using the local information $\mathcal{J}_i = \boldsymbol{x}_i - \boldsymbol{s}_i$ of the agent i, and the received neighboring information set $\mathcal{I}_i = \bigcup_{j\in\mathcal{N}_i} \mathcal{J}_j$ is shared from its neighbors. The reference information $\mathcal{J}_{r,i}$ for each agent i is defined according to the connectivity with the leader as

$$\mathcal{J}_{r,i} = \begin{cases} \{\boldsymbol{x}_l, x_{l,r}, \boldsymbol{s}_i, s_{i,r}\}, & \text{if } b_{ii} = 1 \\ \{\boldsymbol{s}_i, s_{i,r}\}, & \text{if } b_{ii} = 0 \end{cases}. \quad (6)$$

Given the $\mathcal{J}_i$, $\mathcal{I}_i$ and $\mathcal{J}_{r,i}$, the linear consensus law $\phi(\cdot)$ can be evaluated only with the local and neighbor information, allowing distributed computation, which is designed as

$$\phi(\mathcal{J}_i, \mathcal{I}_i, \mathcal{J}_{r,i}) = -cr_i + s_{i,r} + b_{ii}x_{l,r}, \quad (7)$$

where control gain denotes $c \in \mathbb{R}_+$ and the filtered error is $r_i = \sum_{k=1}^n \lambda_k e_{i,k}$. The synchronization errors $e_{i,k}$ for the k-th dimension are calculated as

$$e_{i,k} = \sum\nolimits_{j\in\mathcal{N}_i} a_{ij}(\tilde{x}_{i,k} - \tilde{x}_{j,k}) + b_{ii}\vartheta_{i,k} \quad (8)$$

for $\forall k=1,\cdots,n$ with $\vartheta_{i,k}=x_{i,k}-s_{i,k}-x_{l,k}$ and $\tilde{x}_{i,k}=x_{i,k}-s_{i,k}$. Meanwhile, the coefficients $\lambda_1,\cdots,\lambda_{n-1}\in\mathbb{R}$ and $\lambda_n\in\mathbb{R}_+$ are chosen such that the matrix $\mathbf{\Lambda}\in\mathbb{R}^{(n-1)\times(n-1)}$ is Hurwitz with

$$\mathbf{\Lambda}=\begin{bmatrix}\mathbf{0}_{(n-2)\times 1} & \boldsymbol{I}_{n-2}\\ -\lambda_n^{-1}\lambda_1 & -\lambda_n^{-1}[\lambda_2,\cdots,\lambda_{n-1}]\end{bmatrix}. \tag{9}$$

Remark 2: The control law in (5) is commonly found for high-order MAS (1) aligned with many applications, such as robotics [34], [35] with $n=2$ and electrohydraulic systems [36] with $n=4$. For other system classes such as strict feedback nonlinear systems [37] and non-affine systems [38], [39], nontrivial design of the distributed control law is required, e.g., through a special-shaped Laplacian matrix. Considering the focus of this paper is on the event-trigger design for cooperative online learning, the extension of the derived method to these systems is left in future works.

The compensation functions $\hat{f}_i(\cdot)$ in (5) predict the unknown function $f(\cdot)$ through the data-driven method using the data set collected individually by each agent $i\in\mathcal{V}$, which satisfies the following assumption.

Assumption 3: The data pair $\{\boldsymbol{x}_i^{(\varsigma)},y_i^{(\varsigma)}\}$ is available for each agent $i\in\mathcal{V}$ at any time $t_i^{(\varsigma)}\in\mathbb{R}_{0,+},\varsigma\in\mathbb{N}_+$, where $y_i^{(\varsigma)}=\dot{x}_{i,n}^{(\varsigma)}+w_i^{(\varsigma)}$. The measurement noise $w_i^{(\varsigma)}\in\mathbb{R}$ of $y_i^{(\varsigma)}$ follows a zero-mean, independent and identical Gaussian distribution, i.e., $w_i^{(\varsigma)}\sim\mathcal{N}(0,\sigma_{o,i}^2)$ with $\sigma_{o,i}>0$.

Assumption 3 indicates each agent collects the data set on their own, without sharing among them. The requirement of full state measurement, i.e., $\boldsymbol{x}_i$, is usually found in controller design for nonlinear system [30] and MAS [12], [16]. Moreover, Assumption 3 allows noisy observations of $f(\cdot)$ due to known $g(\cdot)$ and $h(\cdot)$ and through numerical approximation of $\dot{x}_{i,n}^{(\varsigma)}$, e.g., via finite difference inducing Gaussian error [28].

Additionally, Assumption 3 also facilitates the online data collection for the prediction model update. For GP model update with high data efficiency, an event-triggered mechanism is required for online learning in MAS, that employs a smart data selection strategy to store only the necessary data and ensures a desired control performance. To this end, the event-triggered mechanism is designed such that $\{\boldsymbol{x}_i^{(\varsigma)},y_i^{(\varsigma)}\}$ is added into the data set of agent i at time $t_i^{(\varsigma)}$ satisfying

$$t_i^{(\varsigma)}=\inf\{t_i:t_i>t_i^{(\varsigma-1)}\ \wedge\ \rho_i(t_i)>\bar{\rho}_i(t_i)\}, \tag{10}$$

where $\rho_i(t_i)$ and $\bar{\rho}_i(t_i)$ are the simplified versions of the trigger function and its threshold function, respectively. For centralized event-trigger, it has $\rho_i(t_i)=\rho(\cup_{j\in\mathcal{V}}\{\mathcal{J}_j(t_i),\mathcal{J}_{r,j}(t_i)\})$ and $\bar{\rho}_i(t_i)=\bar{\rho}(\cup_{j\in\mathcal{V}}\{\mathcal{J}_j(t_i),\mathcal{J}_{r,j}(t_i)\})$ requiring global information for evaluation. In the distributed scenario, $\rho_i(t_i)$ and $\bar{\rho}_i(t_i)$ only employ local and neighboring information, i.e., $\rho_i(t_i)=\rho(\mathcal{J}_i(t_i),\mathcal{I}_i(t_i),\mathcal{J}_{r,i}(t_i))$ and $\bar{\rho}_i(t_i)=\bar{\rho}(\mathcal{J}_i(t_i),\mathcal{I}_i(t_i),\mathcal{J}_{r,i}(t_i))$, allowing distributed computation on each agent $i\in\mathcal{V}$. Note that we set both $\rho(\cdot)$ and $\bar{\rho}(\cdot)$ as time-varying functions for better explanability of the trigger condition (10), whose detailed design, i.e., explicit expression of $\rho(\cdot)$ and $\bar{\rho}(\cdot)$, for centralized and distributed scenarios are described in Section III and Section IV, respectively.

III. Cooperative Online Learning based Distributed Control with Gaussian Processes

A. Gaussian Process Regression

A Gaussian process induces a distribution of the unknown function $f(\cdot)$ characterized by the mean function $m(\cdot):\mathbb{X}\to\mathbb{R}$ and kernel function $\kappa(\cdot,\cdot):\mathbb{X}\times\mathbb{X}\to\mathbb{R}_{0,+}$, i.e., $f(\cdot)\sim\mathcal{GP}(m(\cdot),\kappa(\cdot,\cdot))$. Particularly, the mean function $m(\cdot)$ reflects the prior knowledge, which is set as $m(\cdot)=0$ considering $f(\cdot)$ is fully unknown, since all known part in the dynamics (1) is encoded into $h(\cdot)$. And the kernel function $\kappa(\cdot,\cdot)$ indicates the covariance between two samples and is assumed to satisfy the following condition.

Assumption 4: The continuous function $f(\cdot)$ belongs to a function space defined by a Gaussian process $\mathcal{GP}(0,\kappa(\cdot,\cdot))$ with stationary and Lipschitz continuous kernel function with Lipschitz constant $L_\kappa\in\mathbb{R}_+$. Moreover, the kernel function $\kappa(\boldsymbol{x},\boldsymbol{x}')=\kappa(\|\boldsymbol{x}-\boldsymbol{x}'\|)$ is monotonically decreasing with respect to $\|\boldsymbol{x}-\boldsymbol{x}'\|$, and $\kappa(0)=\sigma_f^2$ with $\sigma_f\in\mathbb{R}_+$.

Assumption 4 defines the prior distribution of unknown function $f(\cdot)$ by choosing suitable kernel. The Lipschitz continuity of $\kappa(\cdot,\cdot)$ only requires the kernel function to be continuous by considering the compact input domain $\mathbb{X}$, which holds for most kernels, such as square exponential kernel, rational quadratic kernel, and their combination [6]. Therefore, this assumption imposes no significant restrictions.

Given the data set $\mathbb{D}$ with $M\in\mathbb{N}$ samples, i.e., $\mathbb{D}=\{\boldsymbol{x}^{(\varsigma)},y^{(\varsigma)}\}_{\varsigma=1}^M$, satisfying Assumption 3 under the variance of measurement noise as σ_o^2 with $\sigma_o>0$. Then the posterior mean $\mu(\cdot)$ and variance $\sigma^2(\cdot)$ of the GP model are

$$\begin{aligned}\mu(\boldsymbol{x})&=\boldsymbol{k}_X^T(\boldsymbol{x})(\boldsymbol{K}+\sigma_o^2\boldsymbol{I}_M)^{-1}\boldsymbol{y},\\ \sigma^2(\boldsymbol{x})&=\kappa(\boldsymbol{x},\boldsymbol{x})-\boldsymbol{k}_X^T(\boldsymbol{x})(\boldsymbol{K}+\sigma_o^2\boldsymbol{I}_M)^{-1}\boldsymbol{k}_X(\boldsymbol{x}),\end{aligned} \tag{11}$$

where $\boldsymbol{y}=[y^{(1)},\cdots,y^{(M)}]^T$, $\boldsymbol{K}=\{\kappa(\boldsymbol{x}^{(i)},\boldsymbol{x}^{(j)})\}_{i,j=1,\cdots,M}$ and $\boldsymbol{k}_X(\boldsymbol{x})=[\kappa(\boldsymbol{x},\boldsymbol{x}^{(1)}),\cdots,\kappa(\boldsymbol{x},\boldsymbol{x}^{(M)})]^T$. The posterior mean $\mu(\cdot)$ is used for the estimation of $f(\cdot)$, while the posterior variance $\sigma^2(\cdot)$ is employed to quantify the prediction error as shown in the following lemma.

Lemma 2 ([40]): Predict an unknown function $f(\cdot)$ satisfying Assumption 4 using GP regression with a data set satisfying Assumption 3. Pick $\tau\in\mathbb{R}_+$ and $\delta\in(0,1)\subset\mathbb{R}$, the prediction error is upper bounded as

$$|f(\boldsymbol{x})-\mu(\boldsymbol{x})|\le\eta_\delta(\boldsymbol{x})=\sqrt{\beta_\delta}\sigma(\boldsymbol{x})+\gamma_\delta,\ \forall\boldsymbol{x}\in\mathbb{X}$$

with a probability of at least $1-\delta$, where

$$\begin{aligned}\beta_\delta&=2\sum\nolimits_{k=1}^n\log\Big(\frac{\sqrt{n}}{2\tau}(\bar{x}_k-\underline{x}_k)+1\Big)-2\log\delta,\\ \gamma_\delta&=\big(\sqrt{\beta_\delta}L_\sigma+L_f+L_\mu\big)\tau\end{aligned}$$

with $\bar{x}_k=\max_{\boldsymbol{x}\in\mathbb{X}}x_k$ and $\underline{x}_k=\min_{\boldsymbol{x}\in\mathbb{X}}x_k$ for x_k as the k-th dimension of $\boldsymbol{x}$, and L_f, L_μ and L_σ are the Lipschitz constants for the unknown function $f(\cdot)$, the posterior mean $\mu(\cdot)$ and variance $\sigma(\cdot)$, respectively.

Although with conservatism, Lemma 2 provides a calculable uniform prediction error bound on the compact domain $\mathbb{X}$. The detailed computations for Lipschitz constants L_μ, L_σ and L_f refer to [40]. Due to the computable error bound

$\eta_\delta(\cdot)$, GP regression is widely used in safe learning-based control with guarantee [16], [19], [41].

B. Cooperative Online Learning

In our setting, a GP model is deployed on each agent i with individual data set $\mathbb{D}_i$ satisfying Assumption 3 with measurement noise variance $\sigma_{o,i}$ and kernel function $\kappa_i(\cdot)$ under Assumption 4 with $\kappa_i(0)=\sigma_{f,i}^2$ and $\sigma_{f,i}\in\mathbb{R}_+$. To obtain the prediction of $f(\cdot)$ cooperatively in MASs, each agent i shares its states $\boldsymbol{x}_i$ with its bidirectional neighbors in $\bar{\mathcal{N}}_i$, such that its neighbor agents $\forall j\in\bar{\mathcal{N}}_i$ calculate the prediction $\mu_j(\boldsymbol{x}_i)$ at $\boldsymbol{x}_i$ using their own data set $\mathbb{D}_j$. Notably, only the predictions from agents in $\bar{\mathcal{N}}_i$ are available on agent i, since any agent j belonging to $\mathcal{N}_i\backslash\bar{\mathcal{N}}_i$ cannot calculate the prediction $\mu_j(\boldsymbol{x}_i)$ for agent i. This is because $\boldsymbol{x}_i$ cannot send to agent j due to the lack of transmission channel from j to i, i.e., $(i,j)\notin\mathcal{E}$. Combining the local prediction $\mu_i(\boldsymbol{x}_i)$ and the received neighboring predictions $\{\mu_j(\boldsymbol{x}_i)\}_{j\in\bar{\mathcal{N}}_i}$, the compensation $\hat{f}_i(\boldsymbol{x}_i)$ in (5) is formulated using aggregation method with a general form as

$$\hat{f}_i(\boldsymbol{x}_i)=\omega_{ii}(\boldsymbol{x}_i)\mu_i(\boldsymbol{x}_i)+\sum\nolimits_{j\in\bar{\mathcal{N}}_i}\omega_{ij}(\boldsymbol{x}_i)\mu_j(\boldsymbol{x}_i),\tag{12}$$

where $\omega_{ij}(\cdot):\mathbb{X}\to\mathbb{R}_{0,+}$ indicates the aggregation weight respective to $\mu_j(\boldsymbol{x}_i)$ defined for specific aggregation such as POE [13]. Since online learning strategy is applied such that each agent i adds newly generated data pair $\{\boldsymbol{x}_i^{(\varsigma)},y_i^{(\varsigma)}\}$ at $t_i^{(\varsigma)}$ into the local data set $\mathbb{D}_i$. The aggregated prediction error bound after the GP model update is shown as follows.

Lemma 3: Let the assumptions in Lemma 2 be satisfied for all local GP models on the agents in the MAS, and employ the aggregation method in (12) for agent $i\in\mathcal{V}$. Define $\underline{\hat{\sigma}}_i$ as the solution of the optimal problem for $i\in\mathcal{V}$ as

$$\begin{aligned}&\underline{\hat{\sigma}}_i=\sup_{\boldsymbol{x}_i\in\mathbb{X}}\hat{\sigma}_i^+(\boldsymbol{x})=\sup_{\boldsymbol{x}_i\in\mathbb{X}}\Big(\omega_{ii}^+(\boldsymbol{x}_i)\sigma_i^+(\boldsymbol{x}_i)+\sum_{j\in\bar{\mathcal{N}}_i}\omega_{ij}^+(\boldsymbol{x}_i)\sigma_j(\boldsymbol{x}_i)\Big),\\&\text{s.t. }\sigma_i^+(\boldsymbol{x}_i)=\sigma_{o,i}\sigma_i(\boldsymbol{x}_i)/(\sigma_i^2(\boldsymbol{x}_i)+\sigma_{o,i}^2)^{1/2},\end{aligned}\tag{13}$$

where the aggregation weight $\omega_{ij}^+(\boldsymbol{x}_i)$ for agent $j\in\{i,\bar{\mathcal{N}}_i\}$ is evaluated at $\boldsymbol{x}_i$ after the GP model update at agent i. Moreover, let $\omega_{ij}^+(\boldsymbol{x}_i)$ maintain the property of $\sum_{j\in\{i,\bar{\mathcal{N}}_i\}}\omega_{ij}^+(\cdot)=1$ and choose $\delta\in(0,N^{-1})$, then the prediction error after GP model update by using the cooperative learning in (12) is bounded as

$$|f(\boldsymbol{x}_i)-\hat{f}_i^+(\boldsymbol{x}_i)|\le\hat{\eta}_{\delta,i}=\sqrt{\beta_\delta}\underline{\hat{\sigma}}_i+\gamma_\delta,\qquad\forall\boldsymbol{x}_i\in\mathbb{X},$$

with probability of at least $1-N\delta$, where $\hat{f}_i^+(\cdot)$ is the aggregated prediction after local model update at agent i.

Proof: Considering each agent i collects the data pairs on its own and adds them into the local data set $\mathbb{D}_i$, the posterior variance $\sigma_i^+(\boldsymbol{x})$ after GP model update agent i is written as

$$\begin{aligned}(\sigma_i^+(\boldsymbol{x}))^2=&\kappa_i(0)\\&-\begin{bmatrix}\boldsymbol{k}_{X,i}(\boldsymbol{x})\\\kappa_i(0)\end{bmatrix}^T\begin{bmatrix}\boldsymbol{K}_i+\sigma_{o,i}^2\boldsymbol{I}_{M_i}&\boldsymbol{k}_{X,i}(\boldsymbol{x})\\\boldsymbol{k}_{X,i}^T(\boldsymbol{x})&\kappa_i(0)+\sigma_{o,i}^2\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{k}_{X,i}(\boldsymbol{x})\\\kappa_i(0)\end{bmatrix}\\=&\sigma_i^2(\boldsymbol{x})-\frac{\sigma_i^4(\boldsymbol{x})}{\sigma_i^2(\boldsymbol{x})+\sigma_{o,i}^2}=\frac{\sigma_i^2(\boldsymbol{x})\sigma_{o,i}^2}{\sigma_i^2(\boldsymbol{x})+\sigma_{o,i}^2},\end{aligned}\tag{14}$$

where M_i is the number of training samples in $\mathbb{D}_i$ before adding the collected data pairs. The kernel vector $\boldsymbol{k}_{X,i}(\boldsymbol{x})$ and Gram matrix $\boldsymbol{K}_i$ follow the definition in (11) using the kernel function $\kappa_i(\cdot,\cdot)$ evaluated with $\mathbb{D}_i$. Note that (14) is equivalent to the expression of the constraint in (13), such that the prediction error bound at agent i is written as

$$\Pr\{|\mu_i^+(\boldsymbol{x})-f(\boldsymbol{x})|\le\eta_i^+(\boldsymbol{x})=\sqrt{\beta_\delta}\sigma_i^+(\boldsymbol{x})+\gamma_\delta,\forall\boldsymbol{x}\in\mathbb{X}\}\ge1-\delta$$

from Lemma 2, where $\mu_i^+(\cdot)$ denotes the posterior mean after model update on agent i. Recall the condition for the aggregation weights as $\sum_{j\in\{i,\bar{\mathcal{N}}_i\}}\omega_{ij}^+(\cdot)=1$, the aggregated prediction error is bounded by

$$\begin{aligned}|f(\boldsymbol{x}_i)-\hat{f}_i^+(\boldsymbol{x}_i)|\le&\omega_{ii}^+(\boldsymbol{x}_i)|f(\boldsymbol{x}_i)-\mu_i^+(\boldsymbol{x}_i)|\\&+\sum\nolimits_{j\in\bar{\mathcal{N}}_i}\omega_{ij}^+(\boldsymbol{x}_i)|f(\boldsymbol{x}_i)-\mu_j(\boldsymbol{x}_i)|.\end{aligned}\tag{15}$$

Moreover, consider the prediction error bound for predictions from agents $j\in\bar{\mathcal{N}}_i$ according to Lemma 2 as $\Pr\{|f(\boldsymbol{x}_i)-\mu_j(\boldsymbol{x}_i)|\le\eta_j(\boldsymbol{x}_i),\forall\boldsymbol{x}_i\in\mathbb{X}\}\ge1-\delta$, the aggregated prediction error is bounded by

$$\begin{aligned}|f(\boldsymbol{x}_i)-\hat{f}_i^+(\boldsymbol{x}_i)|&\le\omega_{ii}^+(\boldsymbol{x}_i)\eta_i^+(\boldsymbol{x}_i)+\sum\nolimits_{j\in\bar{\mathcal{N}}_i}\omega_{ij}^+(\boldsymbol{x}_i)\eta_j(\boldsymbol{x}_i)\\&=\sqrt{\beta_\delta}\big(\omega_{ii}^+(\boldsymbol{x}_i)\sigma_i^+(\boldsymbol{x}_i)+\sum\nolimits_{j\in\bar{\mathcal{N}}_i}\omega_{ij}^+(\boldsymbol{x}_i)\sigma_j(\boldsymbol{x}_i)\big)+\gamma_\delta\\&=\sqrt{\beta_\delta}\hat{\sigma}_i^+(\boldsymbol{x}_i)+\gamma_\delta\le\sqrt{\beta_\delta}\underline{\hat{\sigma}}_i+\gamma_\delta=\underline{\hat{\eta}}_{\delta,i}\end{aligned}$$

with probability of at least $1-N\delta\le1-|\bar{\mathcal{N}}_i|\delta$ by using the union bound in the first inequality [42]. ■

Lemma 3 shows the aggregated prediction error bound $\underline{\hat{\eta}}_{\delta,i}$ for each agent $i\in\mathcal{V}$ with online learning under the condition of aggregation weights as $\sum_{j\in\{i,\bar{\mathcal{N}}_i\}}\omega_{ij}(\cdot)=1$. This condition is set to overcome the explosive prediction variance when leaving the training data [11], and is commonly found in most well-known aggregation methods, such as mixture of experts (MOE, [43]) with the form

$$\omega_{ij}(\boldsymbol{x}_i)=\omega_{ij},\qquad\forall i\in\mathcal{V},j\in\bar{\mathcal{N}}_i\tag{16}$$

and product of experts [13] with the form

$$\omega_{ij}(\boldsymbol{x}_i)=\frac{\omega_{ij}^*\sigma_j^{-2}(\boldsymbol{x}_i)}{\sum_{k\in\bar{\mathcal{N}}_i}\omega_{ik}^*\sigma_k^{-2}(\boldsymbol{x}_i)},\qquad\forall i\in\mathcal{V},j\in\bar{\mathcal{N}}_i\tag{17}$$

with auxiliary constants $\omega_{ij}^*\in\mathbb{R}_{0,+}$. Note that the choice of ω_{ij}^* for POE can be arbitrary, i.e., not necessary to satisfy $\sum_{j\in\{i,\bar{\mathcal{N}}_i\}}\omega_{ij}^*=1$, while the constants for MOE requires the condition of $\sum_{j\in\{i,\bar{\mathcal{N}}_i\}}\omega_{ij}=1$. With the explicit formulations of the aggregation methods (12) as in (16) and (17), the solution $\underline{\hat{\sigma}}_i$ of (13) has the closed form calculated as follows.

Corollary 1: The solution $\underline{\hat{\sigma}}_i$ in (13) with aggregation methods MOE in (16) and POE in (17) are derived as follows:

(i) For MOE, $\underline{\hat{\sigma}}_i=\omega_{ii}(\sigma_{f,i}^{-2}+\sigma_{o,i}^{-2})^{-\frac{1}{2}}+\sum_{j\in\bar{\mathcal{N}}_i}\omega_{ij}\sigma_{f,j}$; (18)

(ii) For POE, $\underline{\hat{\sigma}}_i^2=\dfrac{\sum_{j\in\{i,\bar{\mathcal{N}}_i\}}\omega_{ij}^*}{\omega_{ii}^*(\sigma_{o,i}^{-2}+\sigma_{f,i}^{-2})+\sum_{j\in\bar{\mathcal{N}}_i}\omega_{ij}^*\sigma_{f,j}^{-2}}$. (19)

Proof: (i) Due to the constant aggregation weights ω_{ij} in (16), the monotonically increasing of $\hat{\sigma}_i$ w.r.t $\hat{\sigma}_i^+(\boldsymbol{x}_i)$ and $\sigma_j(\boldsymbol{x}_i)$ for $j\in\bar{\mathcal{N}}_i$ is obvious. Considering the definition of $\hat{\sigma}_i^+(\boldsymbol{x}_i)$ in (13), it is derived $\hat{\sigma}_i^+(\boldsymbol{x}_i)\le(\sigma_{f,i}^{-2}+\sigma_{o,i}^{-2})^{-1/2}$,

due to the monotonic increasing of $\hat{\sigma}_i^+(\boldsymbol{x}_i)$ w.r.t $\hat{\sigma}_i(\boldsymbol{x}_i)$ with $\hat{\sigma}_i(\boldsymbol{x}_i) \leq \sigma_{f,i}$ for $\forall \boldsymbol{x}_i \in \mathbb{X}$ and positive $\kappa_i(\cdot,\cdot)$. Similarly, considering $\sigma_j(\boldsymbol{x}_i) \leq \sigma_{f,j}$ from (11) with positive definite kernel $\kappa_j(\cdot,\cdot)$, the result in (18) is straightforwardly obtained. (ii) For POE, apply (17) into (12) and consider the definition of $\hat{\sigma}_i^+(\boldsymbol{x}_i)$ in (13), the updated variance $\hat{\sigma}_i^+(\boldsymbol{x}_i)$ is written as

$$\hat{\sigma}_i^+(\boldsymbol{x}_i) = \frac{\omega_{ii}^*(\sigma_i^+(\boldsymbol{x}_i))^{-1} + \sum_{j\in\bar{\mathcal{N}}_i}\omega_{ij}^*\sigma_j^{-1}(\boldsymbol{x}_i)}{\omega_{ii}^*(\sigma_i^+(\boldsymbol{x}_i))^{-2} + \sum_{j\in\bar{\mathcal{N}}_i}\omega_{ij}^*\sigma_j^{-2}(\boldsymbol{x}_i)}, \quad (20)$$

which is not monotonic w.r.t the posterior variance $\sigma_j(\cdot)$ from each agent $j \in \{i, \bar{\mathcal{N}}_i\}$. Then, applying the Cauchy-Schwarz inequality on the numerator of (20), it yields

$$\Big(\sum_{j\in\{i,\bar{\mathcal{N}}_i\}}\omega_{ij}^*\sigma_j^{-1}(\boldsymbol{x}_i)\Big)^2 \leq \Big(\sum_{j\in\{i,\bar{\mathcal{N}}_i\}}\omega_{ij}^*\Big)\Big(\sum_{j\in\{i,\bar{\mathcal{N}}_i\}}\omega_{ij}^*\sigma_j^{-2}(\boldsymbol{x}_i)\Big), \quad (21)$$

such that $\hat{\sigma}_i(\boldsymbol{x}_i)$ in (20) is bounded by

$$\hat{\sigma}_i(\boldsymbol{x}_i) \leq \Big(\sum_{j\in\{i,\bar{\mathcal{N}}_i\}}\omega_{ij}^*\Big)^{1/2}\Big(\omega_{ii}^*(\sigma_j^+(\boldsymbol{x}_i))^{-2} + \sum_{j\in\{i,\bar{\mathcal{N}}_i\}}\omega_{ij}^*\sigma_j^{-2}(\boldsymbol{x}_i)\Big)^{-1/2},$$

where the right-hand side is monotonically increasing w.r.t $\sigma_j(\boldsymbol{x}_i)$ for $j \in \{i, \bar{\mathcal{N}}_i\}$. Therefore, the supremum of the right hand side is achieved when $\hat{\sigma}_i^+(\boldsymbol{x}) = (\sigma_{f,i}^{-2} + \sigma_{o,i}^{-2})^{-1/2}$ and $\sigma_j(\boldsymbol{x}) = \sigma_{f,j}$ for $j \in \bar{\mathcal{N}}_i$, leading to the result in (19). ∎

Corollary 1 shows exact expression of $\underline{\hat{\sigma}}_i$ with the aggregation methods using MOE (16) and POE (17). Note that the prediction error bound of online cooperative learning is non-zero due to the presence of measurement noise $\sigma_{o,i}$ and uncontrollable prediction performance from neighbors $j \in \bar{\mathcal{N}}_i$, resulting in conservative estimation by using $\sigma_{f,j}$.

Moreover, Lemma 3 also allows the derivation of the overall prediction error bound using online cooperative learning, which is shown as follows.

Corollary 2: Let all assumptions in Lemma 3 hold, then the concatenated prediction error for all agents $i \in \mathcal{V}$ is bounded by $\|\hat{\boldsymbol{f}}^+(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})\| \leq \|\underline{\hat{\boldsymbol{\eta}}}_\delta\| = \|[\underline{\hat{\eta}}_{\delta,1}, \cdots, \underline{\hat{\eta}}_{\delta,N}]^T\|$ with probability of at least $1 - N^2\delta$, where $\hat{\boldsymbol{f}}^+(\boldsymbol{x}) = [\hat{f}_1^+(\boldsymbol{x}_1), \cdots, \hat{f}_N^+(\boldsymbol{x}_N)]^T$ and $\boldsymbol{f}(\boldsymbol{x}) = [f(\boldsymbol{x}_1), \cdots, f(\boldsymbol{x}_N)]^T$.

Proof: The overall prediction error is written as $\|\hat{\boldsymbol{f}}^+(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})\|^2 = \sum_{i\in\mathcal{V}}|f(\boldsymbol{x}_i) - \hat{f}_i^+(\boldsymbol{x}_i)|^2 \leq \sum_{i\in\mathcal{V}}\hat{\eta}_{\delta,i}^2 = \|\underline{\hat{\boldsymbol{\eta}}}_\delta\|^2$ with probability of at least $1 - N^2\delta$ due to the union bound, which concludes the proof. ∎

Corollary 2 shows the overall prediction error bound, which is used to determine the control performance, i.e., tracking error bound, in Section III-C.

C. Control Performance Analysis

In this subsection, the control performance, in particular the ultimate boundness of the error $\boldsymbol{\vartheta}$, is analyzed through Lyapunov theorem for the MAS with (1) controlled by the distributed controller (5) with cooperative learning in (12). With the time-varying data sets and GP models due to the online learning, the control law (5) is piece-wise continuous and leads to a switching controller. This results in a closed-loop hybrid system, whose stability is analyzed through common Lyapunov function approach. For a high order MAS in (1), the common Lyapunov candidate V is usually chosen as

$$V = V_1 + V_2 \quad \text{with} \quad V_1 = \boldsymbol{r}^T\boldsymbol{P}_r\boldsymbol{r}, \quad V_2 = \boldsymbol{\varepsilon}^T\boldsymbol{P}_\varepsilon\boldsymbol{\varepsilon}, \quad (22)$$

where $\boldsymbol{\varepsilon} = [\boldsymbol{\varepsilon}_1^T, \cdots, \boldsymbol{\varepsilon}_{n-1}^T]^T \in \mathbb{R}^{(n-1)N}$ is defined as the concatenated synchronization error with $\boldsymbol{\varepsilon}_k = [e_{1,k}, \cdots, e_{N,k}]^T \in \mathbb{R}^N, \forall k = 1, \cdots, n$. The positive definite matrix $\boldsymbol{P}_r$ for $\boldsymbol{r}$ is chosen as $\boldsymbol{P}$ in Lemma 1 such that $\boldsymbol{Q}_r = \boldsymbol{P}_r(\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}}) + (\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}})^T\boldsymbol{P}_r$. And the weight matrix $\boldsymbol{P}_\varepsilon$ for $\boldsymbol{\varepsilon}$ is calculated by $\boldsymbol{P}_\varepsilon = \boldsymbol{P}_{\varepsilon,s} \otimes \boldsymbol{I}_N$, where the positive definite matrix $\boldsymbol{P}_{\varepsilon,s}$ is the solution of the continuous Lyapunov equation $\boldsymbol{\Lambda}^T\boldsymbol{P}_{\varepsilon,s} + \boldsymbol{P}_{\varepsilon,s}\boldsymbol{\Lambda} = -\boldsymbol{Q}_{\varepsilon,s}$ for a given symmetric positive definite $\boldsymbol{Q}_{\varepsilon,s}$. Note that the existence and uniqueness of $\boldsymbol{P}_{\varepsilon,s}$ are guaranteed by considering $\boldsymbol{\Lambda}$ in (9) as Hurwitz. Moreover, define $\boldsymbol{z} = [\boldsymbol{r}^T, \boldsymbol{\varepsilon}^T]^T$ and $\boldsymbol{P}_z = \text{blkdiag}(\boldsymbol{P}_r, \boldsymbol{P}_\varepsilon)$, then the Lyapunov candidate in (22) is reformulated as $V = \boldsymbol{z}^T\boldsymbol{P}_z\boldsymbol{z}$. Owing to the unknown system dynamics $f(\cdot)$ and considering a non-zero prediction error bound as specified in Lemma 3, it is intuitive that the overall tracking error denoted as $\boldsymbol{\vartheta} = \boldsymbol{x} - \boldsymbol{s} - \boldsymbol{1}_N \otimes \boldsymbol{x}_l$ cannot be nullified, instead it is upper bounded by $\bar{\vartheta}$ described in Section II-A. This also leads to a non-zero guaranteed lower bound of V, confines the variable $\boldsymbol{z}$ within certain bounds considering Eq. (22). Specifically, the relationship between $\boldsymbol{z}$ and $\bar{\vartheta}$ is shown in the following lemma.

Lemma 4: Consider the MAS (1) connected through a communication network satisfying Assumption 2 and controlled by (5). Choose the Lyapunov candidate as (22) and

$$\chi = \|(\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}})^{-1}\|((1 + \|[\boldsymbol{t}, \boldsymbol{\Lambda}]\|^2)\underline{\lambda}^{-1}(\boldsymbol{P}_z)\bar{\lambda}(\boldsymbol{P}_z))^{1/2}. \quad (23)$$

If there exists a positive constant $\bar{z} \in \mathbb{R}_{0,+}$, such that the negativity of $\dot{V}$ is shown for $\forall\|\boldsymbol{z}\| > \bar{z}$. Choose $\bar{\vartheta} = \chi\bar{z}$, then the tracking error is ultimately bounded by $\bar{\vartheta}$, i.e., $\|\boldsymbol{\vartheta}\| \leq \bar{\vartheta}$.

Proof: Considering the identical leader for each agent, the synchronization error in (8) is reformulated as $e_{i,k} = \sum_{j\in\mathcal{N}_i} a_{ij}(\vartheta_{i,k} - \vartheta_{j,k}) + b_{ii}\vartheta_{i,k}$, such that their concatenations $\boldsymbol{\varepsilon}_k$ for all dimensions $k \in 1, \cdots, n-1$ are written as

$$\boldsymbol{\varepsilon}_k = (\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}})\tilde{\boldsymbol{\vartheta}}_k, \quad [\boldsymbol{\varepsilon}^T, \boldsymbol{\varepsilon}_n^T]^T = (\boldsymbol{I}_n \otimes (\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}}))\tilde{\boldsymbol{\vartheta}} \quad (24)$$

with $\tilde{\boldsymbol{\vartheta}} = [\tilde{\boldsymbol{\vartheta}}_1^T, \cdots, \tilde{\boldsymbol{\vartheta}}_n^T]^T$, $\tilde{\boldsymbol{\vartheta}}_k = [\vartheta_{1,k}, \cdots, \vartheta_{N,k}]^T$ and $\boldsymbol{\varepsilon}_n = [e_{1,n}, \cdots, e_{N,n}]^T$. Since $\tilde{\boldsymbol{\vartheta}}$ and $\boldsymbol{\vartheta}$ are both composed of $\vartheta_{i,k}$ inducing $\|\tilde{\boldsymbol{\vartheta}}\| = \|\boldsymbol{\vartheta}\|$, the norm of $\boldsymbol{\vartheta}$ is bounded using (24) as

$$\|\boldsymbol{\vartheta}\| = \|\tilde{\boldsymbol{\vartheta}}\| \leq \|(\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}})^{-1}\|\|[\boldsymbol{\varepsilon}_1^T, \dot{\boldsymbol{\varepsilon}}^T]\|, \quad (25)$$

considering $\|[\boldsymbol{\varepsilon}^T, \boldsymbol{\varepsilon}_n^T]^T\| = \|[\boldsymbol{\varepsilon}_1^T, \dot{\boldsymbol{\varepsilon}}^T]\|$. Combining the definition of the filtered error $\boldsymbol{r}$ in (22) and synchronization error $e_{i,k}$ in (8), the derivative of $\boldsymbol{\varepsilon}$ denotes

$$\dot{\boldsymbol{\varepsilon}} = (\boldsymbol{\Lambda} \otimes \boldsymbol{I}_N)\boldsymbol{\varepsilon} + (\boldsymbol{t} \otimes \boldsymbol{I}_N)\boldsymbol{r} = ([\boldsymbol{t}, \boldsymbol{\Lambda}] \otimes \boldsymbol{I}_N)\boldsymbol{z} \quad (26)$$

with $\boldsymbol{t} = [0, \cdots, 0, \lambda_n^{-1}]^T \in \mathbb{R}^{n-1}$, such that the overall tracking error in (25) is further bounded by

$$\begin{aligned}\|\boldsymbol{\vartheta}\|^2 &\leq \|(\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}})^{-1}\|^2(\|\boldsymbol{\varepsilon}_1\|^2 + \|\dot{\boldsymbol{\varepsilon}}\|^2) \\ &\leq \|(\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}})^{-1}\|^2(1 + \|[\boldsymbol{t}, \boldsymbol{\Lambda}]\|^2)\|\boldsymbol{z}\|^2\end{aligned}$$

considering $\|\boldsymbol{\varepsilon}_1\| \leq \|\boldsymbol{z}\|$, since $\boldsymbol{\varepsilon}_1$ is only one part of $\boldsymbol{z}$. Furthermore, considering that V is in quadratic form as in (22) inducing $V \geq \underline{\lambda}(\boldsymbol{P}_z)\|\boldsymbol{z}\|^2$, it has

$$\|\boldsymbol{\vartheta}\| \leq \|(\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}})^{-1}\|((1 + \|[\boldsymbol{t}, \boldsymbol{\Lambda}]\|^2)\underline{\lambda}^{-1}(\boldsymbol{P}_z)V)^{1/2}. \quad (27)$$

Next, we show the upper bound of V. Considering the case when $V > \bar{\lambda}(\boldsymbol{P}_z)\chi^{-2}\bar{\vartheta}^2$, where $\|\boldsymbol{z}\| > \chi^{-1}\bar{\vartheta} = \bar{z}$ is directly derived due to the fact that $V \leq \bar{\lambda}(\boldsymbol{P}_z)\|\boldsymbol{z}\|^2$. Moreover, according to the setting in the lemma, $\|\boldsymbol{z}\| > \bar{z}$ leads to $\dot{V} < 0$, i.e., the decrease of V. Therefore, V is ultimately bounded as $V \leq \bar{\lambda}(\boldsymbol{P}_z)\chi^{-2}\bar{\vartheta}^2$. Apply the boundness of V into (27), then the result in the lemma is straightforwardly derived. ■

Lemma 4 shows the overall tracking error $\boldsymbol{\vartheta}$ is bounded, if the positive value $\bar{z}$ exists and is well defined such that $\dot{V} < 0$ when $\|\boldsymbol{z}\| > \bar{z}$. The expression of $\bar{z}$ and then $\bar{\vartheta}$ is derived by observing the time derivative of V in (22) and using Lemma 4, whose result is shown in the following theorem.

Theorem 1: Consider the control of MAS (1) with Assumption 1 under a communicate graph satisfying Assumption 2. The compensation $\hat{f}_i^+(\cdot)$ is obtained by aggregation online GP predictions satisfying Assumption 3 and $\sum_{j\in\{i,\mathcal{N}_i\}}\omega_{ij}(\cdot) = 1$. Pick $\delta \in (0, 1/N^2)$ and choose $c \in \mathbb{R}_+$ in (5), such that $\boldsymbol{Q}_z \in \mathbb{R}^{nN\times nN}$ is positive definite defined as

$$\boldsymbol{Q}_z = \begin{bmatrix} c\lambda_n\boldsymbol{Q}_r - 2\lambda_n\lambda_{n-1}^{-1}\boldsymbol{P}_r & -\boldsymbol{\Psi} \\ -\boldsymbol{\Psi}^T & \boldsymbol{Q}_\varepsilon \end{bmatrix} \tag{28}$$

with $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_{n-1}]^T$, $\boldsymbol{\Psi} = \boldsymbol{P}_r(\boldsymbol{\lambda}^T\boldsymbol{\Lambda}\otimes\boldsymbol{I}_N) + (\boldsymbol{t}^T\boldsymbol{P}_{\varepsilon,s})\otimes\boldsymbol{I}_N$ and $\boldsymbol{Q}_\varepsilon = \boldsymbol{Q}_{\varepsilon,s}\otimes\boldsymbol{I}_N$. Then, the tracking error $\|\boldsymbol{\vartheta}\|$ is bounded by $\bar{\vartheta} = \chi\xi\|\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta\|$ with a probability of at least $1 - N^2\delta$, where $\boldsymbol{\iota} = [\iota_1, \cdots, \iota_N]^T \in \mathbb{R}^N$, $\iota_i = (1-b_{ii})F_l$, $\forall i = 1, \cdots, N$ with F_l defined in Assumption 1 and

$$\xi = 2\lambda_n\underline{\lambda}^{-1}(\boldsymbol{Q}_z)\|\boldsymbol{P}_r(\boldsymbol{\mathcal{L}}+\boldsymbol{\mathcal{B}})\|. \tag{29}$$

Proof: The tracking error bound $\bar{\vartheta}$ is derived through common Lyapunov theory, where the sign of $\dot{V} = \dot{V}_1 + \dot{V}_2$ from (22) is investigated. First, we show the time derivative of V_1 in (22), which is written as $\dot{V}_1 = \dot{\boldsymbol{r}}^T\boldsymbol{P}_r\boldsymbol{r} + \boldsymbol{r}^T\boldsymbol{P}_r\dot{\boldsymbol{r}}$, where

$$\dot{\boldsymbol{r}} = \sum\nolimits_{k=1}^{n-1}\lambda_k\dot{\boldsymbol{\varepsilon}}_k + \lambda_n\dot{\boldsymbol{\varepsilon}}_n = (\boldsymbol{\lambda}^T\otimes\boldsymbol{I}_N)\dot{\boldsymbol{\varepsilon}} + \lambda_n\dot{\boldsymbol{\varepsilon}}_n.$$

Considering the definition of $\boldsymbol{\varepsilon}_n$ with the synchronization errors in (8) and the system dynamics (1), $\dot{\boldsymbol{\varepsilon}}_n$ is formulated as

$$\dot{\boldsymbol{\varepsilon}}_n = (\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}})(\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{G}(\boldsymbol{x})\boldsymbol{u} - \boldsymbol{s}_r - \boldsymbol{1}_N x_{l,r}), \tag{30}$$

where $\boldsymbol{f}(\boldsymbol{x}) = [f(\boldsymbol{x}_1), \cdots, f(\boldsymbol{x}_N)]^T$. Applying the consensus law $\phi(\cdot)$ in (7) and compensation $\hat{\boldsymbol{f}}^+(\cdot)$ into (5), the concatenated control input $\boldsymbol{u} = [u_1, \cdots, u_N]^T$ is written as

$$\boldsymbol{u} = \boldsymbol{G}^{-1}(\boldsymbol{x})(-c\boldsymbol{r} + \boldsymbol{s}_r + \boldsymbol{\mathcal{B}}\boldsymbol{1}_N x_{l,r} - \hat{\boldsymbol{f}}^+(\boldsymbol{x})), \tag{31}$$

where $\boldsymbol{r} = [r_1, \cdots, r_N]^T$, $\boldsymbol{s}_r = [s_{r,1}, \cdots, s_{r,N}]^T$ and $\boldsymbol{G}(\boldsymbol{x}) = \mathrm{diag}(g(\boldsymbol{x}_1), \cdots, g(\boldsymbol{x}_N))$. Moreover, taking $\boldsymbol{u}$ from (31) and $\dot{\boldsymbol{\varepsilon}}$ from (26) into (30), the derivative of $\boldsymbol{r}$ is written as $\dot{\boldsymbol{r}} = (\boldsymbol{\lambda}^T\boldsymbol{t}\otimes\boldsymbol{I}_N - c\lambda_n(\boldsymbol{\mathcal{L}}+\boldsymbol{\mathcal{B}}))\boldsymbol{r} + (\boldsymbol{\lambda}^T\boldsymbol{\Lambda}\otimes\boldsymbol{I}_N)\boldsymbol{\varepsilon} + \lambda_n(\boldsymbol{\mathcal{L}}+\boldsymbol{\mathcal{B}})\boldsymbol{\psi}^+$, where $\boldsymbol{\psi}^+ = (\boldsymbol{\mathcal{B}} - \boldsymbol{I}_N)\boldsymbol{1}_N x_{l,r} + \boldsymbol{f}(\boldsymbol{x}) - \hat{\boldsymbol{f}}^+(\boldsymbol{x})$. Due to the definition of $\boldsymbol{Q}_r$ and the fact that $\boldsymbol{\lambda}^T\boldsymbol{t} = \lambda_n\lambda_{n-1}^{-1}$, it has

$$\begin{aligned}\dot{V}_1 = &-\boldsymbol{r}^T(c\lambda_n\boldsymbol{Q}_r - 2\lambda_n\lambda_{n-1}^{-1}\boldsymbol{P}_r)\boldsymbol{r} \\ &+ 2\boldsymbol{r}^T\boldsymbol{P}_r\big((\boldsymbol{\lambda}^T\boldsymbol{\Lambda}\otimes\boldsymbol{I}_N)\boldsymbol{\varepsilon} + \lambda_n(\boldsymbol{\mathcal{L}}+\boldsymbol{\mathcal{B}})\boldsymbol{\psi}^+\big).\end{aligned} \tag{32}$$

Moreover, considering $\dot{\boldsymbol{\varepsilon}}$ in (26) and Lemma 1, the derivative of V_2 in (22) is straightforwardly derived as follows

$$\dot{V}_2 = -\boldsymbol{\varepsilon}^T\boldsymbol{Q}_\varepsilon\boldsymbol{\varepsilon} + 2\boldsymbol{\varepsilon}^T(\boldsymbol{P}_{\varepsilon,s}\boldsymbol{t}\otimes\boldsymbol{I}_N)\boldsymbol{r}. \tag{33}$$

Combining with (32) and (33) and using the definition of $\boldsymbol{z}$, the derivative of V is written as

$$\begin{aligned}\dot{V} &= -\boldsymbol{z}^T\boldsymbol{Q}_z\boldsymbol{z} + 2\lambda_n\boldsymbol{r}^T\boldsymbol{P}_r(\boldsymbol{\mathcal{L}}+\boldsymbol{\mathcal{B}})\boldsymbol{\psi}^+ \\ &\leq -\underline{\lambda}(\boldsymbol{Q}_z)\|\boldsymbol{z}\|^2 + 2\lambda_n\|\boldsymbol{z}\|\|\boldsymbol{P}_r(\boldsymbol{\mathcal{L}}+\boldsymbol{\mathcal{B}})\|\|\boldsymbol{\psi}^+\| \\ &\leq -\underline{\lambda}(\boldsymbol{Q}_z)\|\boldsymbol{z}\|(\|\boldsymbol{z}\| - \xi\|\boldsymbol{\psi}^+\|),\end{aligned} \tag{34}$$

where the first and second inequalities are satisfied due to the positive definite of $\boldsymbol{Q}_z$ from the choice of c in (28) and the definition of ξ in (29), respectively. Because of the uncertainty of unknown function $f(\cdot)$, it is not possible to compute the exact norm of $\boldsymbol{\psi}^+$. Instead, the upper bound of $\|\boldsymbol{\psi}^+\|$ can be derived using both the triangular inequality and the prediction error bound described in Corollary 2. Specifically, the norm of $\boldsymbol{\psi}^+$ is bounded under Assumption 1 and $b_{ii} \leq 1, \forall i \in \mathcal{V}$ as

$$\begin{aligned}\|\boldsymbol{\psi}^+\| &\leq \|(\boldsymbol{I}_N - \boldsymbol{\mathcal{B}})\boldsymbol{1}_N|x_{l,r}| + |\boldsymbol{f}(\boldsymbol{x}) - \hat{\boldsymbol{f}}^+(\boldsymbol{x})|\| \\ &\leq \|(\boldsymbol{I}_N - \boldsymbol{\mathcal{B}})\boldsymbol{1}_N F_l + \hat{\boldsymbol{\eta}}_\delta^+(\boldsymbol{x})\| = \|\boldsymbol{\iota} + \hat{\boldsymbol{\eta}}_\delta^+(\boldsymbol{x})\|\end{aligned} \tag{35}$$

with a probability of at least $1 - N^2\delta$ using the union bound, where $\hat{\boldsymbol{\eta}}_\delta^+(\boldsymbol{x}) = [\hat{\eta}_{\delta,1}^+(\boldsymbol{x}_1), \cdots, \hat{\eta}_{\delta,N}^+(\boldsymbol{x}_N)]^T$ with $\hat{\eta}_{\delta,i}^+(\boldsymbol{x}_i) = \sqrt{\beta_\delta}\sigma_i^+(\boldsymbol{x}_i) + \gamma_\delta$ and $\sigma_i^+(\boldsymbol{x}_i)$ defined in Lemma 3. The vector $\boldsymbol{\iota} = (\boldsymbol{I}_N - \boldsymbol{\mathcal{B}})\boldsymbol{1}_N F_l$ with positive entries denotes the leader misconnection term with the bounded $|x_{l,r}|$ by F_l from Assumption 1. Moreover, considering the prediction error bound of the cooperative online learning in Lemma 3, the norm of $\boldsymbol{\psi}^+$ in (35) is further bounded by $\|\boldsymbol{\psi}^+\| \leq \|\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta\|$, leading to the probabilistic bound as

$$\dot{V} \leq -\underline{\lambda}(\boldsymbol{Q}_z)\|\boldsymbol{z}\|(\|\boldsymbol{z}\| - \xi\|\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta\|). \tag{36}$$

Defining $\bar{z}$ in Lemma 4 as $\bar{z} = \xi\|\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta\|$, the negativity of $\dot{V}$ is achieved when $\|\boldsymbol{z}\| > \bar{z}$ from (36). Then, apply the result in Lemma 4, the boundness of $\|\boldsymbol{\vartheta}\|$ in Theorem 1 is derived, inheriting the probability of at least $1 - N^2\delta$. ■

Theorem 1 shows the tracking error bound $\bar{\vartheta}$ under the distributed control with cooperative online learning strategy. The tracking error bound $\bar{\vartheta}$ also reflected by $\bar{z}$ is related to both the prediction accuracy and connectivity with the leader reflected by $\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta$ with additional coefficient ξ. Despite the non-zero $\underline{\hat{\boldsymbol{\eta}}}_\delta$ from Corollary 2 and non-zero $\boldsymbol{\iota}$ from leader misconnection, the arbitrary small tracking error bound $\bar{\vartheta}$ can be achieved by choosing sufficiently large control gain c such that $\underline{\lambda}(\boldsymbol{Q}_z)$ is large inducing small ξ and thus small $\bar{z}$.

However, Theorem 1 requires $\hat{\eta}_{\delta,i}(\cdot)$ to be smaller than $\underline{\hat{\eta}}_{\delta,i}$ for each agent $i \in \mathcal{V}$ without specifying the data collection strategy. In the next section, smart data selection strategies are proposed, aiming to enhance data efficiency while maintaining the desired control performance.

IV. Event-triggered Online Learning

This section delves into the development of intelligent online data selection strategies for efficient training data storage, leveraging event-triggered mechanisms. It commences with a centralized event-triggered approach in Section IV-A, followed by the distributed event-trigger in Section IV-B, accompanied by the theoretical performance guarantee. Furthermore, in both centralized and distributed event-triggered cooperative online learning, exclusion of Zeno behavior is shown in Section IV-C.

Note that with event-triggered online learning mechanism, only GP models on some of agents will be updated at time t. For notational simplicity, define a time related index function $\varpi_i(\cdot):\mathbb{R}_{0,+}\to\{0,1\}$ for agent $i\in\mathcal{V}$ as

$$\varpi_i(t)=\begin{cases}1, & \text{if agent } i \text{ is updated at } t\\ 0, & \text{otherwise}\end{cases}. \tag{37}$$

With the index function $\varpi_i(\cdot)$, the aggregated prediction error bound for agent i is defined as $\tilde{\eta}_{\delta,i}(\cdot)$with expression

$$\tilde{\eta}_{\delta,i}(\boldsymbol{x}_i(t))=\varpi_i(t)\hat{\eta}^+_{\delta,i}(\boldsymbol{x}_i(t))+(1-\varpi_i(t))\hat{\eta}_{\delta,i}(\boldsymbol{x}_i(t)) \tag{38}$$

for all $i\in\mathcal{V}$, where $\hat{\eta}_{\delta,i}(\cdot)$ and $\hat{\eta}^+_{\delta,i}(\cdot)$ are recalled as the error bound with and without online learning, respectively. Moreover, the concatenated prediction error bound denotes $\tilde{\boldsymbol{\eta}}_\delta(\boldsymbol{x})=[\tilde{\eta}_{\delta,1}(\boldsymbol{x}_1),\cdots,\tilde{\eta}_{\delta,N}(\boldsymbol{x}_N)]^T$.

A. Centralized Event-triggered Online Learning

Before the design of distributed even-trigger online learning, the centralized version is first analyzed here, which assumes the existence of a centralized node collecting global information and then transmitting the online learning decision to each relevant agent. The centralized event-trigger for cooperative learning is devised by extending the previous research for single agent in [19], [28], [29] to multi-agent setting. Specifically to maintain a predefined upper tracking error bound $\bar{\vartheta}_c>\bar{\vartheta}$ with $\bar{\vartheta}$ defined in Theorem 1, the centralized trigger function and its threshold are designed as

$$\rho=\|\boldsymbol{\iota}+\hat{\boldsymbol{\eta}}_\delta(\boldsymbol{x})\|,\qquad \bar{\rho}=\xi^{-1}\max\{\|\boldsymbol{z}\|,\chi^{-1}\bar{\vartheta}_c\}, \tag{39}$$

where the time input in (39) is dropped, i.e., $\rho(t)\to\rho$, for notational simplicity under the assumption of all the employed variables in the same time instances. Note that the condition $\bar{\vartheta}_c>\bar{\vartheta}$ indicates there exist constants $\epsilon_i\in\mathbb{R}_+$ selected for each agent $i\in\mathcal{V}$, such that $\bar{\vartheta}_c$ is reformulated as $\bar{\vartheta}_c=\xi\chi\|\boldsymbol{\iota}+\underline{\hat{\boldsymbol{\eta}}}_\delta+\boldsymbol{\epsilon}\|$ with $\boldsymbol{\epsilon}=[\epsilon_1,\cdots,\epsilon_N]^T$. The strictly positive ϵ_i excludes the Zeno behavior on agent i, which is shown in Section IV-C. The control performance is shown as follows.

Proposition 1: Let all assumptions in Theorem 1 hold. Choose $\epsilon_i\in\mathbb{R}_+$ for each agent $i\in\mathcal{V}$ and adopt the online learning strategy with the centralized event-trigger (39) for $\bar{\vartheta}_c=\xi\chi\|\boldsymbol{\iota}+\underline{\hat{\boldsymbol{\eta}}}_\delta+\boldsymbol{\epsilon}\|$. Design $c_i(\cdot):\{0,1\}\to\mathbb{R}$ as the "cost" for model update at agent $i\in\mathcal{V}$. If $\rho(t)>\bar{\rho}(t)$, update the GP models with $\varpi_i=1$ by adding new samples into the local data set, where ϖ_i for $i\in\mathcal{V}$ are obtained by solving

$$\min_{\{\varpi_i\}_{i\in\mathcal{V}}}\sum\nolimits_{i\in\mathcal{V}}c_i(\varpi_i),\quad \text{s.t.}\|\boldsymbol{\iota}+\tilde{\boldsymbol{\eta}}_\delta(\boldsymbol{x})+\boldsymbol{\epsilon}\|\le\|\boldsymbol{\iota}+\underline{\hat{\boldsymbol{\eta}}}_\delta+\boldsymbol{\epsilon}\| \tag{40}$$

with each entry of $\tilde{\boldsymbol{\eta}}_\delta(\boldsymbol{x})$ defined in (37) and (38) respectively. Pick $\delta\in(0,N^{-2})$, then the tracking error $\|\boldsymbol{\vartheta}\|$ is bounded by $\bar{\vartheta}_c$ with probability of at least $1-N^2\delta$.

Proof: See appendix. ∎

Remark 3: To satisfy the same constraint of (40), there exists a heuristic selection method, which only update the local GP model with $\hat{\eta}_{\delta,i}(\boldsymbol{x}_i)\le\underline{\hat{\eta}}_{\delta,i}$, such that the trigger condition for each agent $i\in\mathcal{V}$ becomes

$$\rho_i=\rho\big(\hat{\eta}_{\delta,i}(\boldsymbol{x}_i)-\underline{\hat{\eta}}_{\delta,i}\big),\quad \bar{\rho}_i=\max\{\bar{\rho}\big(\hat{\eta}_{\delta,i}(\boldsymbol{x}_i)-\underline{\hat{\eta}}_{\delta,i}\big),0\}. \tag{41}$$

This method indicates agent i will update its local GP only at t when both $\hat{\eta}_{\delta,i}(\boldsymbol{x}_i)>\underline{\hat{\eta}}_{\delta,i}$ and $\rho(t)>\bar{\rho}(t)$ are satisfied.

Note that the condition in (39) is equivalent to choosing $\rho^*=\|\boldsymbol{\iota}+\hat{\boldsymbol{\eta}}_\delta(\boldsymbol{x})\|^2$ and $\bar{\rho}^*=\xi^{-2}\max\{\|\boldsymbol{z}\|^2,\chi^{-2}\bar{\vartheta}_c^2\}$, where the trigger threshold $\bar{\rho}^*$ is also written as

$$\bar{\rho}^*=\xi^{-2}\max\{\|\boldsymbol{z}\|^2-\chi^{-2}\bar{\vartheta}_c^2,0\}+\|\boldsymbol{\iota}+\underline{\hat{\boldsymbol{\eta}}}_\delta+\boldsymbol{\epsilon}\|^2. \tag{42}$$

The form in (42) is regarded as the combination of transient requirement with $\boldsymbol{z}$, desired tracking error bound $\bar{\vartheta}_c$ and guaranteed performance after model update with $\underline{\hat{\boldsymbol{\eta}}}_\delta$. Intuitively, (42) can be converted to the distributed version by decomposition of $\boldsymbol{z}$ and considering the individual prediction performance guarantee after model update, which is detailed discussed in the next subsection.

B. Distributed Event-triggered Online Learning

For robustness and scalability of the information topology, the distributed event-triggered mechanism is proposed inspired from (42). To achieve the tracking error bound $\bar{\vartheta}_d=\xi\chi\|\boldsymbol{\iota}+\underline{\hat{\boldsymbol{\eta}}}_\delta+\boldsymbol{\epsilon}\|$ with positive entries of $\boldsymbol{\epsilon}$ as $\epsilon_i,\forall i\in\mathcal{V}$, the distributed trigger condition in (10) is written for each agent $i\in\mathcal{V}$ as

$$\begin{aligned}\rho_i&=\big(\hat{\eta}_{\delta,i}(\boldsymbol{x}_i)+\iota_i\big)^2\\ \bar{\rho}_i&=\xi^{-2}\max\big\{\|\boldsymbol{z}_i\|^2-\chi^{-2}\bar{\vartheta}_d^2/N,0\big\}+\big(\underline{\hat{\eta}}_{\delta,i}+\iota_i+\epsilon_i\big)^2,\end{aligned} \tag{43}$$

where $\|\boldsymbol{z}_i\|=[r_i,e_{i,1},\cdots,e_{i,n-1}]^T$, $\iota_i=(1-b_{ii})F_l$ recalled from (35). Similarly to (39), the time input is dropped, i.e., $\rho_i(t)\to\rho_i$ and $\bar{\rho}_i(t)\to\bar{\rho}_i$, for notational simplicity. Note that all terms in ρ_i and $\bar{\rho}_i$ in (43) including $\boldsymbol{z}_i$ can be calculated using only local and neighboring information, realizing distributed evaluation.

Considering the maximum operator in (43) and comparing the values of ρ_i and $\bar{\rho}_i$, the agents are divided into 4 cases based on the criterion as

- Safe set $\mathbb{S}$: Agent $i\in\mathbb{S}$ means $\|\boldsymbol{z}_i\|\le\chi^{-1}\bar{\vartheta}_d/\sqrt{N}$;
- Trigger set $\mathbb{T}$: Agent $i\in\mathbb{T}$ means trigger condition with (43) is satisfied, i.e., $\rho_i>\bar{\rho}_i$.

Moreover, define $\bar{\mathbb{S}}$, $\bar{\mathbb{T}}$ as the complement sets for $\mathbb{S}$, $\mathbb{T}$ respectively, satisfying $\mathbb{S}\cup\bar{\mathbb{S}}=\mathbb{T}\cup\bar{\mathbb{T}}=\mathcal{V}$ and $\mathbb{S}\cap\bar{\mathbb{S}}=\mathbb{T}\cap\bar{\mathbb{T}}=\emptyset$. Note that there exists overlap between $\mathbb{S}$ and $\mathbb{T}$, such that the 4 cases for agents are defined as $\mathbb{S}\cap\mathbb{T}$, $\bar{\mathbb{S}}\cap\mathbb{T}$, $\mathbb{S}\cap\bar{\mathbb{T}}$ and $\bar{\mathbb{S}}\cap\bar{\mathbb{T}}$ with the following properties.

Property 1: (i) If $i\in\mathbb{T}$ or $i\in\mathbb{S}\cap\bar{\mathbb{T}}$, then $\tilde{\eta}_{\delta,i}(\boldsymbol{x}_i)\le\underline{\hat{\eta}}_{\delta,i}+\epsilon_i$;
(ii) If $i\in\bar{\mathbb{S}}\cap\mathbb{T}$, then $\|\boldsymbol{z}_i\|>\chi^{-1}\bar{\vartheta}_d/\sqrt{N}$;
(iii) If $i\in\bar{\mathbb{S}}\cap\bar{\mathbb{T}}$, then $\|\boldsymbol{z}_i\|^2\ge\chi^{-2}\bar{\vartheta}_d^2/N+\xi^2(\rho_i-(\underline{\eta}_{\delta,i}+\iota_i+\epsilon_i)^2)$.

Proof: See appendix. ∎

Property 1 shows relevant properties of the distributed event-trigger in (43) for stability analysis. Considering (36), where $\|\boldsymbol{z}\|$ and $\hat{\boldsymbol{\eta}}_\delta(\cdot)$ contribute the negative and positive part in $\dot{V}$, Property 1 ensures the positive term is sufficiently small when $i\in\mathbb{S}$ by observing Property 1 (i), while the negative term is sufficiently large when $i\in\bar{\mathbb{S}}$ from Property 1 (ii & iii).

The control performance with distributed event-triggered online cooperative learning (43) is shown as follows.

Theorem 2: Consider the MAS (1), where agents communicate through a network satisfying Assumption 2. The unknown function $f(\cdot)$ is predicted by using GP regression

satisfying Assumption 3 and 4 and the aggregation method (12). The control task is to track the reference trajectory satisfying Assumption 1 and $\boldsymbol{x}_i(0) = \boldsymbol{s}_i(0) + \boldsymbol{x}_l(0)$, $\forall i \in \mathcal{V}$ at initial time $t = 0$. For such a task, employ the proposed control law (5) with (7) and the distributed event-trigger mechanism for online learning with (43) and $\epsilon_i \in \mathbb{R}_+, \forall i \in \mathcal{V}$. Pick $\delta \in (0, N^{-2})$, and then the tracking error $\|\boldsymbol{\vartheta}\|$ is bounded by $\bar{\vartheta}_d = \xi\chi\|\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta + \boldsymbol{\epsilon}\|$ with probability of at least $1 - N^2\delta$.

Proof: Given a desired tracking error bound $\bar{\vartheta}$ and using the inverse result in Lemma 4, it requires to prove $\dot{V} < 0$ if $\|\boldsymbol{z}\| > \chi^{-1}\bar{\vartheta}$. Dividing the agents according to the maximum operator and trigger condition into $\mathbb{S} \cap \mathbb{T}$, $\bar{\mathbb{S}} \cap \mathbb{T}$, $\mathbb{S} \cap \bar{\mathbb{T}}$ and $\bar{\mathbb{S}} \cap \bar{\mathbb{T}}$ with properties shown in Property 1, the concatenated effects of the prediction error and leader misconnection, i.e., $\boldsymbol{\iota} + \tilde{\boldsymbol{\eta}}_\delta(\boldsymbol{x}) + \boldsymbol{\epsilon}$, is upper bounded by

$$\begin{aligned}\|\boldsymbol{\iota} + \tilde{\boldsymbol{\eta}}_\delta(\boldsymbol{x}) + \boldsymbol{\epsilon}\|^2 &= \sum_{i\in\mathbb{T}}(\iota_i + \hat{\eta}^+_{\delta,i}(\boldsymbol{x}_i) + \epsilon_i)^2 + \sum_{i\in\mathbb{S}\cap\bar{\mathbb{T}}}\rho_i + \sum_{i\in\bar{\mathbb{S}}\cap\bar{\mathbb{T}}}\rho_i \\ &\le \sum\nolimits_{i\in\mathbb{T}}(\iota_i + \underline{\hat{\eta}}_{\delta,i} + \epsilon_i)^2 + \sum\nolimits_{i\in\mathbb{S}\cap\bar{\mathbb{T}}}(\iota_i + \underline{\hat{\eta}}_{\delta,i} + \epsilon_i)^2 + \sum\nolimits_{i\in\bar{\mathbb{S}}\cap\bar{\mathbb{T}}}\rho_i \\ &= \sum\nolimits_{i\in\mathcal{V}}(\iota_i + \underline{\hat{\eta}}_{\delta,i} + \epsilon_i)^2 + \sum\nolimits_{i\in\bar{\mathbb{S}}\cap\bar{\mathbb{T}}}(\rho_i - (\iota_i + \underline{\hat{\eta}}_{\delta,i} + \epsilon_i)^2).\end{aligned}$$

Moreover, considering $\sum_{i\in\mathcal{V}}(\iota_i + \underline{\tilde{\eta}}_{\delta,i} + \epsilon_i)^2 = \|\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta + \boldsymbol{\epsilon}\|^2 = (\xi\chi)^{-2}\bar{\vartheta}_d^2/N$ from the condition of $\bar{\vartheta}_d$ in Theorem 2, the upper bound for $\|\boldsymbol{\iota} + \tilde{\boldsymbol{\eta}}_\delta(\boldsymbol{x}) + \boldsymbol{\epsilon}\|^2$ is further written as

$$\begin{aligned}\|\boldsymbol{\iota} + \hat{\boldsymbol{\eta}}^+_\delta(\boldsymbol{x}) + \boldsymbol{\epsilon}\|^2 <& (\xi\chi)^{-2}\bar{\vartheta}_d^2/N \\ &+ \sum\nolimits_{i\in\bar{\mathbb{S}}\cap\bar{\mathbb{T}}}(\rho_i - (\iota_i + \underline{\tilde{\eta}}_{\delta,i} + \epsilon_i)^2).\end{aligned} \tag{44}$$

Furthermore, considering Property 1 the concatenated error $\boldsymbol{z}$ is lower bounded by

$$\begin{aligned}\|\boldsymbol{z}\|^2 &= \sum\nolimits_{i\in\mathcal{V}}\|\boldsymbol{z}_i\|^2 \ge \sum\nolimits_{i\in\bar{\mathbb{S}}\cap\mathbb{T}}\|\boldsymbol{z}_i\|^2 + \sum\nolimits_{i\in\bar{\mathbb{S}}\cap\bar{\mathbb{T}}}\|\boldsymbol{z}_i\|^2 \\ &> \sum_{i\in\bar{\mathbb{S}}\cap\mathbb{T}}\chi^{-2}\bar{\vartheta}_d^2/N + \sum_{i\in\bar{\mathbb{S}}\cap\bar{\mathbb{T}}}(\chi^{-2}\bar{\vartheta}_d^2/N + \xi^2(\rho_i - (\underline{\eta}_{\delta,i} + \iota_i + \epsilon_i)^2)) \\ &= \chi^{-2}\bar{\vartheta}_d^2|\bar{\mathbb{S}}|/N + \sum\nolimits_{i\in\bar{\mathbb{S}}\cap\bar{\mathbb{T}}}\xi^2(\rho_i - (\underline{\eta}_{\delta,i} + \iota_i + \epsilon_i)^2).\end{aligned} \tag{45}$$

Apply (44) and (45) into (36), then $\dot{V}$ is upper bounded by

$$\begin{aligned}\dot{V} &\le -\frac{\underline{\lambda}(\boldsymbol{Q}_z)\|\boldsymbol{z}\|}{\|\boldsymbol{z}\| + \xi\|\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta\|}\left(\|\boldsymbol{z}\|^2 - \xi^2\|\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta\|^2\right) \\ &< -\frac{\underline{\lambda}(\boldsymbol{Q}_z)\|\boldsymbol{z}\|\bar{\vartheta}^2}{N\chi^2(\|\boldsymbol{z}\| + \xi\|\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta\|)}\left(|\bar{\mathbb{S}}| - 1\right).\end{aligned} \tag{46}$$

Next, we recall the case with $\|\boldsymbol{z}\| > \chi^{-1}\bar{\vartheta}_d$, which indicates at least one agent satisfying $\|\boldsymbol{z}_i\| > \chi^{-1}\bar{\vartheta}_d/\sqrt{N}$ by considering the contradiction with $\|\boldsymbol{z}_i\| \le \chi^{-1}\bar{\vartheta}_d/\sqrt{N}$, $\forall i \in \mathcal{V}$ inducing $\|\boldsymbol{z}\|^2 \le \sum_{i\in\mathcal{V}}\chi^{-2}\bar{\vartheta}_d^2/N = (\chi^{-1}\bar{\vartheta}_d)^2$. This also means $|\bar{\mathbb{S}}| \ge 1$ if $\|\boldsymbol{z}\| > \chi^{-1}\bar{\vartheta}_d$. Then, the negativity of $\dot{V}$ in (46) is directly obtained for $\|\boldsymbol{z}\| > \chi^{-1}\bar{\vartheta}_d$. Let $\bar{z} = \chi^{-1}\bar{\vartheta}_d$ and apply the result in Lemma 4, the boundness of $\|\boldsymbol{\vartheta}\|$ by $\bar{\vartheta}_d$ is derived. ■

Theorem 2 shows the bounded tracking error $\|\boldsymbol{\vartheta}\|$ with the proposed distributed event-trigger (43). The stability analysis in Theorem 2 provides a guideline for devising the distributed event-trigger, which means any design satisfying Property 1 guarantee the tracking error bound with $\bar{\vartheta}_d$.

C. Zeno Behavior for Event-triggered Cooperative Learning

The Zeno behavior is an essential problem for event-triggered strategies [44], meaning infinite triggers in finite time. In this subsection, we discuss the Zeno behavior on each agent i using the proposed centralized and distributed event-trigger, i.e., (39) and (43) respectively. Note that the exclusion of the Zeno behavior on each agent $i \in \mathcal{V}$ prohibits the jump of $\boldsymbol{x}_i$, which means it requires limited changing rate of the agent states, i.e., $\dot{\boldsymbol{x}}_i$, under the proposed controller (5). The upper bound of $\dot{\boldsymbol{x}}_i$ for $\forall i \in \mathcal{V}$ is shown in the following lemma.

Lemma 5: Let all the assumptions in Lemma 3 with centralized event-trigger mechanism with (39) and any update model selection methods, or Theorem 2 with distributed event-trigger strategy in (43) hold. Choose $\epsilon_i \in \mathbb{R}_+$ for each agent $i \in \mathcal{V}$, such that $\bar{\vartheta}_e = \xi\chi\|\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta + \boldsymbol{\epsilon}\|$. Pick $\delta \in (0, N^{-2})$, then $\dot{\boldsymbol{x}}_i(t)$, $\forall i \in \mathcal{V}$ is probabilistic bounded, i.e., $\Pr\{|\dot{\boldsymbol{x}}_i(t)| \le F_i, \forall t \in \mathbb{R}_{0,+}\} \ge 1 - N^2\delta$ for $F_i \in \mathbb{R}_+$ defined as

$$F_i = F_{r,i} + \left(1 + \sqrt{2}c(l_{ii} + b_{ii})\lambda^*\right)\bar{\vartheta}_e + (1 - b_{ii})F_l + \bar{\tilde{\eta}}_{\delta,i}, \tag{47}$$

where $\lambda^* = \max_{k=1,\cdots,n}\lambda_k$ and $\bar{\tilde{\eta}}_{\delta,i}$ is obtained by solving

$$\bar{\tilde{\eta}}_{\delta,i} = \sup_{\boldsymbol{x}_i\in\mathbb{X}}\left(\omega_{ii}(\boldsymbol{x}_i)\sigma_i(\boldsymbol{x}_i) + \sum\nolimits_{j\in\bar{\mathcal{N}}_i}\omega_{ij}(\boldsymbol{x}_i)\sigma_j(\boldsymbol{x}_i)\right). \tag{48}$$

The real positive constants $F_{r,i}$ and F_l are the upper bounds for $\|\dot{\boldsymbol{x}}_l + \dot{\boldsymbol{s}}_i\|$ and $|x_{l,r}|$ respectively as in Assumption 1.

Proof: The bound of $\|\dot{\boldsymbol{x}}_i\|$ is derived by finding the supremum of the tracking error $\|\dot{\boldsymbol{\vartheta}}_i\|$ for all $i \in \mathcal{V}$. Apply the control law (5) with (7) and (12), then the error dynamics w.r.t $\boldsymbol{\vartheta}_i = [\vartheta_{i,1}, \cdots, \vartheta_{i,n}]^T$ with $\vartheta_{i,k} = x_{i,k} - s_{i,k} - x_{l,k}$ for the controlled system yields $\dot{\vartheta}_{i,k} = \vartheta_{i,k+1}$, $\forall k = 1, \cdots, n-1$ and

$$\dot{\vartheta}_{i,n} = -cr_i - (1 - b_{ii})x_{l,r} + f(\boldsymbol{x}_i) - \tilde{f}_i(\boldsymbol{x}_i),$$

where $\tilde{f}_i(\cdot) = \varpi_i\hat{f}^+_i(\cdot) + (1 - \varpi_i)\hat{f}_i(\cdot)$ represents the applied aggregated prediction under event-trigger mechanism using the notation ϖ_i in Section IV. Then, $\|\dot{\boldsymbol{\vartheta}}_i\| = [\dot{\vartheta}_{i,1}, \cdots, \dot{\vartheta}_{i,n}]^T$ is bounded using the triangular inequality by

$$\begin{aligned}\|\dot{\boldsymbol{\vartheta}}_i\| &\le \|[\dot{\vartheta}_{i,1}, \cdots, \dot{\vartheta}_{i,n-1}]^T\| + |\dot{\vartheta}_{i,n}| \\ &\le \|[\vartheta_{i,2}, \cdots, \vartheta_{i,n}]^T\| + c|r_i| + (1 - b_{ii})|x_{l,r}| + |f(\boldsymbol{x}_i) - \tilde{f}_i(\boldsymbol{x}_i)|.\end{aligned} \tag{49}$$

Consider Assumption 1 and recall $\tilde{\eta}_{\delta,i}(\cdot)$ in Section IV as the probabilistic prediction error bound of $|f(\cdot) - \tilde{f}_i(\cdot)|$ with probability of at least $1 - N\delta$ inherited from Lemma 3, (49) is reformulated by using $\|[\vartheta_{i,2}, \cdots, \vartheta_{i,n}]^T\| \le \|\boldsymbol{\vartheta}_i\| \le \|\boldsymbol{\vartheta}\|$ as

$$\|\dot{\boldsymbol{\vartheta}}_i\| \le \|\boldsymbol{\vartheta}\| + c|r_i| + (1 - b_{ii})F_l + \tilde{\eta}_{\delta,i}(\boldsymbol{x}_i), \tag{50}$$

which holds with probability of at least $1 - N\delta$. Due to the definition of the filtered error r_i in (7), the norm of r_i is bounded by using Cauchy-Schwarz inequality and considering non-negative λ_k as $|r_i|^2 \le \sum_{k=1}^n \lambda_k^2|e_{i,k}|^2$, where the synchronization errors $|e_{i,k}|$ from (8) is bounded by

$$\begin{aligned}|e_{i,k}|^2 &= \left|\left(b_{ii} + \sum\nolimits_{j\in\mathcal{N}_i}a_{ij}\right)\vartheta_{i,k} + \sum\nolimits_{j\in\mathcal{N}_i}a_{ij}\vartheta_{j,k}\right|^2 \\ &\le \sum\nolimits_{j\in\mathcal{N}_i}a_{ij}^2\vartheta_{j,k}^2 + 2(b_{ii} + l_{ii})\left(\sum\nolimits_{j\in\mathcal{N}_i}a_{ij}|\vartheta_{i,k}||\vartheta_{j,k}|\right) \\ &\quad + (b_{ii} + l_{ii})^2\vartheta_{i,k}^2,\end{aligned} \tag{51}$$

where l_{ii} is recalled as the i-th entry of the diagonal of Laplacian matrix $\boldsymbol{\mathcal{L}}$ such that $l_{ii} \ge a_{ij}, \forall j \in \mathcal{V}$ due to the

non-negative of a_{ij}. Moreover, considering $b_{ii} \geq 0$, it has $0 \leq a_{ij} \leq b_{ii} + l_{ii}, \forall j \in \mathcal{V}$. Using the Young's inequality on $|\vartheta_{i,k}||\vartheta_{j,k}|$, (51) is further bounded by

$$\begin{aligned}|e_{i,k}|^2 \leq & (b_{ii}+l_{ii})^2 \sum\nolimits_{j\in\{i,\mathcal{N}_i\}} \vartheta_{jk}^2 + (b_{ii}+l_{ii})\Big(\sum\nolimits_{j\in\mathcal{N}_i} a_{ij}(\vartheta_{ik}^2+\vartheta_{jk}^2)\Big) \\ \leq & 2(b_{ii}+l_{ii})^2 \sum\nolimits_{j\in\mathcal{V}} \vartheta_{j,k}^2, \end{aligned} \tag{52}$$

considering $\{i, \mathcal{N}_i\} \subseteq \mathcal{V}$ and $a_{ij} = 0, \forall j \notin \mathcal{N}_i$. Apply the boundness of $|e_{i,k}|$ in (52), the upper bound of the filtered error r_i is written as

$$\begin{aligned}|r_i|^2 \leq & \, 2(l_{ii}+b_{ii})^2 \sum\nolimits_{k=1}^{n} \lambda_k^2 \sum\nolimits_{j\in\mathcal{V}} \vartheta_{j,k}^2 \\ \leq & \, 2(l_{ii}+b_{ii})^2 (\lambda^*)^2 \sum\nolimits_{k=1}^{n} \sum\nolimits_{j\in\mathcal{V}} \vartheta_{j,k}^2. \end{aligned}$$

Note that $\|\boldsymbol{\vartheta}\|^2 = \sum_{k=1}^{n}\sum_{j=1}^{N} |\vartheta_{j,k}|^2$, then it is derived that $|r_i| \leq \sqrt{2}(l_{ii}+b_{ii})\lambda^*\|\boldsymbol{\vartheta}\|$. Next, we investigate the upper bound of $\tilde{\eta}_{\delta,i}(\cdot)$ in (50). Since the optimization problem (48) for agent i covers the case in (13) by removing the constraint on $\sigma_i(\cdot)$, it has $\bar{\hat{\eta}}_{\delta,i} \geq \underline{\hat{\eta}}_{\delta,i}$ such that

$$\tilde{\eta}_{\delta,i}(\boldsymbol{x}_i) \leq \varpi \underline{\hat{\eta}}_{\delta,i} + (1-\varpi)\bar{\hat{\eta}}_{\delta,i} \leq \bar{\hat{\eta}}_{\delta,i}, \ \forall \boldsymbol{x}_i \in \mathbb{X}. \tag{53}$$

Apply the upper bound for $|r_i|$ and (53) into (50), then $\|\dot{\boldsymbol{\vartheta}}_i\|$ is upper bounded by $\|\dot{\boldsymbol{\vartheta}}_i\| \leq \big(1+\sqrt{2}c(l_{ii}+b_{ii})\lambda^*\big)\|\boldsymbol{\vartheta}\| + (1-b_{ii})F_l + \bar{\hat{\eta}}_{\delta,i}$. Moreover, considering $\|\dot{\boldsymbol{x}}_i\| = \|\dot{\boldsymbol{x}}_l + \dot{\boldsymbol{s}}_i + \dot{\boldsymbol{\vartheta}}_i\| \leq \|\dot{\boldsymbol{x}}_l + \dot{\boldsymbol{s}}_i\| + \|\dot{\boldsymbol{\vartheta}}_i\|$ and the boundness of the reference in Assumption 1 as well as the tracking error bound $\bar{\vartheta}_e$ from Proposition 1 or Theorem 2, the upper bound F_i in (47) of state changing rate for agent i, i.e., $\|\dot{\boldsymbol{x}}_i\|$, is derived. Furthermore, the probability as at least $1 - N^2\delta$ is inherited by considering the usage of all the local predictions in $\bar{\vartheta}_e$. ■

Besides the bounded state changing rate, the exclusion of the Zeno behavior also requires the continuity of the prediction performance reflected by its error bound $\hat{\eta}_{\delta,i}(\cdot)$ w.r.t. $\boldsymbol{x}_i$, which means the derivative of $\hat{\eta}_{\delta,i}(\boldsymbol{x}_i)$, i.e., $\nabla\hat{\eta}_{\delta,i}(\boldsymbol{x}_i) = \mathrm{d}\hat{\eta}_{\delta,i}(\boldsymbol{x}_i)/\mathrm{d}\boldsymbol{x}_i$, is bounded. However, considering the event-triggered online learning, $\nabla\hat{\eta}_{\delta,i}(\boldsymbol{x}_i(t_i^{(\varsigma)}))$ is not continuous at $t_i^{(\varsigma)}$ due to the change of the GP model with updated data set. Therefore, we intend to bound $\nabla\hat{\eta}_{\delta,i}(\boldsymbol{x}_i(t))$ on agent i in the time interval without model update locally or on the neighbors, i.e., $t_i \in [\underline{t}_i, \bar{t}_i)$, where $\bar{t}$ and $\underline{t}$ are defined as

$$\underline{t}_i = \max\nolimits_{j\in\{i,\bar{\mathcal{N}}_i\},\varsigma\in\mathbb{N}_+} \Big\{ t_j^{(\varsigma)} \in \mathbb{R}_{0,+} : t_j^{(\varsigma)} \leq t_i \Big\}, \tag{54}$$

$$\bar{t}_i = \min\nolimits_{j\in\{i,\bar{\mathcal{N}}_i\},\varsigma\in\mathbb{N}_+} \Big\{ t_j^{(\varsigma)} \in \mathbb{R}_{0,+} : t_j^{(\varsigma)} > t_i \Big\}. \tag{55}$$

Note that considering the definition of $\hat{\eta}_{\delta,i}(\cdot)$ as $\hat{\eta}_{\delta,i}(\cdot) = \sqrt{\beta_\delta}\hat{\sigma}_i(\cdot) + \gamma_\delta$, the boundness of $\nabla\hat{\eta}_{\delta,i}(\boldsymbol{x}_i)$ only requires the bounded $\nabla\hat{\sigma}_i(\boldsymbol{x}_i)$. The existence of a well-defined upper bound for $\|\nabla\hat{\sigma}_i(\boldsymbol{x}_i)\|$ is easy to see by assuming the aggregation weight $\omega_{ij}(\boldsymbol{x}_i), \forall j \in \{i, \bar{\mathcal{N}}_i\}$ is continuous. Then, $\omega_{ij}(\boldsymbol{x}_i)$ is Lipschitz continuous within the compact domain $\mathbb{X}$ with the Lipschitz constant defined as $L_{\omega,ij} \in \mathbb{R}_{0,+}$ for $j \in \{i, \bar{\mathcal{N}}_i\}$ and $i \in \mathcal{V}$. The detailed expression for $L_{\omega,ij}$ is shown in [45]. Moreover, the individual posterior variance $\sigma_i(\cdot)$ is also Lipschitz continuous with Lipschitz constant $L_{\sigma,i}$ as shown in Lemma 2 under Assumption 4, whose detailed expression in [40], [42]. With the Lipschitz aggregation weights and individual posterior variances and the choice of $\omega_{ij} \in \mathbb{R}_{0,+}$ such that $\sum_{j\in\{i,\bar{\mathcal{N}}_i\}} \omega_{ij}(\cdot) = 1$, the upper bound of $\|\nabla\hat{\sigma}_i(\boldsymbol{x}_i)\|$ is obtained for $\forall i \in \mathcal{V}$ by considering

$$\begin{aligned}\|\nabla\hat{\sigma}_i(\boldsymbol{x}_i)\| \leq & \sum_{j\in\{i,\bar{\mathcal{N}}_i\}} \big(\|\nabla\omega_{ij}(\boldsymbol{x}_i)\|\sigma_j(\boldsymbol{x}_i) + \omega_{ij}(\boldsymbol{x}_i)\|\nabla\sigma_j(\boldsymbol{x}_i)\|\big) \\ \leq & \sum\nolimits_{j\in\{i,\bar{\mathcal{N}}_i\}} \big(L_{\omega,ij}\sigma_j(\boldsymbol{x}_i) + \omega_{ij}(\boldsymbol{x}_i)L_{\sigma,j}\big) \\ \leq & \sum\nolimits_{j\in\{i,\bar{\mathcal{N}}_i\}} \big(L_{\omega,ij}\sigma_{f,j} + L_{\sigma,j}\big) \end{aligned}$$

for all $\boldsymbol{x}_i \in \mathbb{X}$, where the third inequality is derived by considering $\sigma_j(\cdot) \leq \sigma_{f,j}$ from (11) and $\omega_{ij}(\cdot) \leq 1$. For notational simplicity, define the upper bound for $\|\nabla\hat{\sigma}_i(\boldsymbol{x}_i)\|$, i.e., Lipschitz constant for $\hat{\sigma}_i(\boldsymbol{x}_i)$, as $\hat{L}_{\sigma,i} = \sum_{j\in\{i,\bar{\mathcal{N}}_i\}}(L_{\omega,ij}\sigma_{f,j} + L_{\sigma,j})$. The boundness in whole time domain $\mathbb{R}_{0,+}$ is directly derived by considering $\|\nabla\hat{\sigma}_i(\cdot)\| \leq \hat{L}_{\sigma,i}$ in $[\underline{t}_i, \bar{t}_i)$ with (54) for $\forall t_i \in \mathbb{R}_{0,+}$.

With the bounded state changing rate as in Lemma 5 and bounded change of prediction performance analyzed above, the exclusion of the Zeno behavior for the both proposed centralized and distributed event-trigger mechanism shown in (39) and (43) is proven as follows.

Proposition 2: Let all assumptions in Lemma 5 hold and use the event-triggered online cooperative learning with either centralized mechanism in (39) with heuristic update model selection in (41) or distributed mechanism in (43), in which $\epsilon_i \in \mathbb{R}_+$ is chosen for each agent $i \in \mathcal{V}$. Pick $\delta \in (0, N^{-2})$, then the inter-event time $\Delta_i^{(\varsigma)} = t_i^{(\varsigma+1)} - t_i^{(\varsigma)}, \forall \varsigma \in \mathbb{N}$ for each agent $i \in \mathcal{V}$ is lower bounded by $\underline{\Delta}_i \in \mathbb{R}_+$ as $\underline{\Delta}_i = (\sqrt{\beta_\delta}F_i\hat{L}_{\sigma,i})^{-1}\epsilon_i$ with probability of at least $1 - N^2\delta$.

Proof: From the design of trigger condition in (42) and (43) with the heuristic model selection strategy in (41), it is obvious the model update will be activated on agent i at $t_i^{(\varsigma+1)}$ if $\rho_i(t_i^{(\varsigma+1)}) > \bar{\rho}_i(t_i^{(\varsigma+1)})$ indicating $\hat{\eta}_{\delta,i}(\boldsymbol{x}_i(t_i^{(\varsigma+1)})) > \underline{\hat{\eta}}_{\delta,i} + \epsilon_i$. Moreover, considering the model update on the agent i occurs at $t_i^{(\varsigma)}$ such that the aggregated prediction error bound after model update is upper bounded as $\hat{\eta}_{\delta,i}^{+}(\boldsymbol{x}_i(t_i^{(\varsigma)})) \leq \underline{\hat{\eta}}_{\delta,i}$ from Lemma 3. Note that no new data pairs are added into the data set $\mathbb{D}_i$ in $(t_i^{(\varsigma)}, t_i^{(\varsigma+1)})$, such that the prediction model on agent i maintains unchanged. With the same prediction model, the difference between $\hat{\eta}_{\delta,i}^{+}(\boldsymbol{x}_i(t_i^{(\varsigma)}))$ and $\hat{\eta}_{\delta,i}(\boldsymbol{x}_i(t_i^{(\varsigma+1)}))$ is lower bounded by ϵ_i, which is also written as

$$\begin{aligned}\epsilon_i < & \, \hat{\eta}_{\delta,i}(\boldsymbol{x}_i(t_i^{(\varsigma+1)})) - \hat{\eta}_{\delta,i}^{+}(\boldsymbol{x}_i(t_i^{(\varsigma)})) \\ \leq & \int_{t_i^{(\varsigma)}}^{t_i^{(\varsigma+1)}} |\dot{\hat{\eta}}_i(\boldsymbol{x}_i(\tau))| d\tau = \sqrt{\beta_\delta} \int_{t_i^{(\varsigma)}}^{t_i^{(\varsigma+1)}} |\dot{\hat{\sigma}}_i(\boldsymbol{x}_i(\tau))| d\tau. \end{aligned} \tag{56}$$

Using the chain rule, the derivative of $\hat{\sigma}_i(\boldsymbol{x}_i)$ is bounded by

$$|\dot{\hat{\sigma}}_i(\boldsymbol{x}_i)| = \Big|\dot{\boldsymbol{x}}_i^T \frac{\mathrm{d}\hat{\sigma}_i(\boldsymbol{x}_i)}{\mathrm{d}\boldsymbol{x}_i}\Big| \leq \|\dot{\boldsymbol{x}}_i\| \Big\| \frac{\mathrm{d}\hat{\sigma}_i(\boldsymbol{x}_i)}{\mathrm{d}\boldsymbol{x}_i} \Big\| \leq F_i\hat{L}_{\sigma,i}$$

with probability of at least $1 - N^2\delta$ using F_i and $\hat{L}_{\sigma,i}$ from Lemma 5. Then, (56) is reformulated as $\epsilon_i < \sqrt{\beta_\delta}F_i\hat{L}_{\sigma,i}\Delta_i^{(\varsigma)}$, indicating the boundness as $\Delta_i^{(\varsigma)} \geq \underline{\Delta}_i$ holds for $\forall \varsigma \in \mathbb{N}$. ■

Proposition 2 shows the minimal trigger interval for each agent is non-zero by using both centralized and distributed event-triggered mechanism for cooperative online learning, indicating the exclusion of the Zeno behavior. Note that although

the proof for centralized event-trigger only works for heuristic update model selection strategy in (41), the Zeno behavior exclusion for other selection methods is easily derived by using the similar methods in [19], [29], which considers the minimal trigger interval for the whole MAS. Moreover, Proposition 2 directly shows, only choosing strict positive ϵ_i leads to guaranteed non-zero minimal trigger interval $\underline{\Delta}_i$, resulting in Zeno behavior exclusion. Combining with the tracking error bound in Proposition 1 and Theorem 2, there exists a trade-off between the control performance and update frequency, i.e., larger ϵ_i induces worse tracking performance but allowing low model update rate.

V. Numerical Simulations

A. Toy Example

In this section[1], we consider a multi-agent system including $N = 4$ agents. Each agent $i \in \mathcal{V} = \{1, 2, 3, 4\}$ follows the dynamics in (1) with $n = 2$, $h(\boldsymbol{x}_i) = 0$, $g(\boldsymbol{x}_i) = 1$ and $f(\boldsymbol{x}_i) = 5\sin(10x_{i,1}) + 0.5/(1 + \exp(x_{i,2}/10)) + 10$, where the input domain is set as $\mathbb{X} = [-1.5, 1.5] \times [-0.6, 0.6]$. The agents are connected with a directed communication network defined by the edge set as $\mathcal{E} = \{(2,3),(3,1),(3,2),(3,4),(4,1),(4,2),(4,3)\}$ and $a_{ij} = 1, \forall (j,i) \in \mathcal{E}$. Moreover, only agent 1 and 3 have the access to the leader, i.e., $b_{ii} = 1$ for $i = 1, 3$ and $b_{ii} = 0$ for $i = 2, 4$. The prediction of $f(\cdot)$ is obtained by (12) using POE [13], such that the aggregation weights $\omega_{ij}(\cdot)$ are calculated as $\omega_{ij}(\boldsymbol{x}_i) = \sigma_j^{-2}(\boldsymbol{x}_i)(\sum_{k\in\bar{\mathcal{N}}_i}\sigma_k^{-2}(\boldsymbol{x}_i))^{-1}$, $\forall i \in \mathcal{V}, j \in \bar{\mathcal{N}}_i$ and $\omega_{ij}(\boldsymbol{x}_i) = 0$ otherwise. Moreover, consider $\hat{\sigma}_i^2 = (|\bar{\mathcal{N}}_i| + 1)((\sigma_{o,i}^{-2} + \sigma_{f,i}^{-2}) + \sum_{j\in\bar{\mathcal{N}}_i}\sigma_{f,j}^{-2})^{-1}$ as the solution of (13), such that it is easy to see $\hat{L}_{\sigma,i} = (|\bar{\mathcal{N}}_i| + 1)^{\frac{1}{2}}\sum_{j\in\{i,\bar{\mathcal{N}}_i\}} L_{\sigma,j}$. The kernel functions $\kappa(\cdot,\cdot)$ in all local Gaussian processes are identical, which is in square exponential form as $\kappa_i(\boldsymbol{x}_i, \boldsymbol{x}_i') = \sigma_{f,i}^2\exp(-0.5l_i^{-2}\|\boldsymbol{x}_i - \boldsymbol{x}_i'\|^2)$, $\forall i \in \mathcal{V}$ with $\sigma_{f,i} = 1, l_i = 0.1$. The data pairs for the prediction satisfy Assumption 3 with the measurements y_i perturbed by Gaussian noise with $\sigma_{o,i} = 0.01$, $\forall i \in \mathcal{V}$. The initial data sets vary depending on the applied learning methods, and are discussed later. The grid factor and probability for uniform error bound in Lemma 2 are set as $\tau = 10^{-6}$ and $\delta = 0.001$, respectively. The control objective is to track the leader trajectory with $x_{l,1} = \sin(2t/5)$, $x_{l,2} = \dot{x}_{l,1}$, $x_{l,r} = \dot{x}_{l,2}$ and maintain relative states for each agent i as $s_{i,1} = 0.01\sin(6t + 2\pi i/N)$, $s_{i,2} = \dot{s}_{i,1}$, $s_{i,r} = \dot{s}_{i,2}$. The control law (5) is designed for each agent i, where the control gains in (7) are set as $c = 20$ and $\lambda_1 = \lambda_2 = 1$. Moreover, select $\boldsymbol{Q}_\varepsilon = 1$ and then the matrix $\boldsymbol{Q}_z$ in (28) is ensured to be positive definite. The simulation time is set to 40.

To demonstrate the effectiveness of the proposed event-triggered online learning method, we compare the tracking error $\|\boldsymbol{\vartheta}(t)\|$ by using the following learning strategies:

1) Offline cooperative learning [12]: Each agent $i \in \mathcal{V}$ provides the prediction $\sigma_i(\cdot)$ using the initial data set, containing 200 random samples uniformly distributed in $\mathbb{X}$;
2) Centralized event-trigger (CET): The trigger in (39) with heuristic update model selection in (41).

[1]The code is available at https://drive.google.com/drive/folders/1cOdToV_h__VWfHNKmi-ib-VioJSpKCGq?usp=sharing.

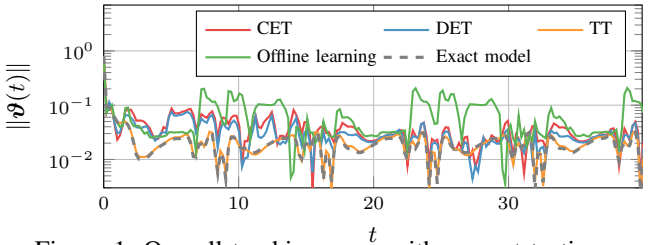


Figure 1: Overall tracking error with respect to time.

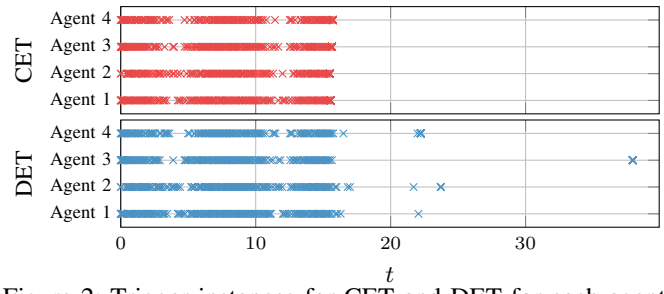


Figure 2: Trigger instances for CET and DET for each agent. Specifically, for CET 131, 130, 122 and 120 samples are collected for agent 1 to 4. The minimal trigger interval for each agent denotes 0.036, 0.027, 0.020 and 0.020. With DET, each agent collects 154, 150, 132 and 131 data pairs under minimal trigger interval 0.025, 0.024, 0.016 and 0.021.

3) Distributed event-trigger (DET): The trigger in (43).
4) Time-trigger (TT) [20]: Update the local GPs on each agent every 0.015 time interval.
5) Exact model: Let $\hat{f}_i(\cdot) = f(\cdot)$ in (5) for $\forall i \in \mathcal{V}$.

Note that the trigger interval for TT, i.e., 0.015, is set such that it is similar to the minimal trigger interval by using CET and DET as shown in Fig. 2. Moreover, for both CET and DET, the initial data sets for each agent are set as empty, and ϵ_i in (15) are chosen as $\epsilon_i = \sigma_{o,i} - (\sigma_{f,i}^{-2} + \sigma_{o,i}^{-2})^{-1/2}$, $\forall i \in \mathcal{V}$. The effectiveness of the above learning methods is shown in the following subsections.

1) Performance with Event-triggered Online Learning: First, we observe the effectiveness of different cooperative learning strategies for MAS with same initial condition, i.e., $\boldsymbol{x}_i(0) = \boldsymbol{x}_l(0) + \boldsymbol{s}_i(0), \forall i \in \mathcal{V}$. The overall tracking error $\|\boldsymbol{\vartheta}(t)\|$ w.r.t time is shown in Fig. 1. It is easy to see, online learning behaviors much better with smaller tracking error compared to offline learning. While with longer transition phase, both event-triggered learning induces similar performance after $t = 20$ as TT. Moreover, the performance between CET and DET is similar, indicating almost no performance loss from the distributed computation.

Besides the tracking error, the number of triggers reflect the data efficiency, where we only consider the triggers for event-triggered online learning mechanisms shown in Fig. 2. It is obvious that most trigger occurs before $t = 18$ due to the periodic references, indicating the collected data set around $t = 18$ is sufficient for the guaranteed tracking performance in Lemma 3. Only few triggers happens after $t = 20$ for DET due to its slight conservatism, which results in near 20% more triggers compared to the CET.

2) Monte Carlo Test: To test the generalization of the proposed event-triggered cooperative online learning methods,

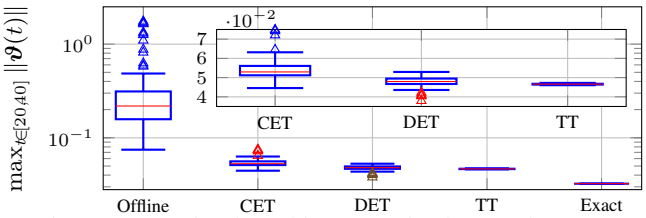


Figure 3: Maximal tracking error in the steady state.

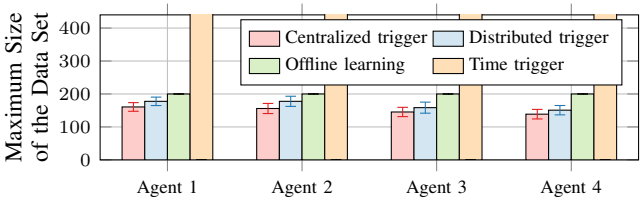


Figure 4: Maximal size of data set, which for offline learning is 200 from the size of initial data set and for TT is 2667 due to the fixed trigger interval 0.015. With CET, the maximal size of the data set for each agent 1 to 4 denotes 157 ± 13, 152 ± 13, 145 ± 14 and 135 ± 12, respectively. The maximal numbers of training samples collected through DET are 177 ± 16, 175 ± 15, 158 ± 17 and 153 ± 15 for agent 1 to 4, respectively.

a Mont-Carlo test is employed, in which each algorithm is repeated for 100 times by using different initial states, i.e., uniformly distributed $\boldsymbol{x}_i(0)\in\mathbb{X}$, uniformly distributed initial data set for offline learning and normally distributed random measurement noise w_i^{ς} for each agent $i\in\mathcal{V}$.

The control performance for each learning algorithm is reflected by the ultimate tracking error bound as shown in Fig. 3. In practice, we consider the maximal tracking error in $t\in[20,40]$, in order to neglect larger $\|\boldsymbol{\vartheta}\|$ from poor initial states $\boldsymbol{x}(0)$ and focus on the steady states. It is obvious, online learning induces much smaller tracking error than offline learning, due to better adaption of the data set for the control task. The slightly worse performance from event-triggered learning compared to TT is due to the underestimation of the performance of GP update by merely using one-point GP accuracy in Lemma 3. Moreover, to evaluate the data efficiency by using event-triggered data collection, the trigger times from each methods are compared.

The number of triggers reflecting the data efficiency is depicted in Fig. 4 with the maximal number of training samples in each data set. It is obvious that both event-triggered online learning mechanisms results in the data set with slightly less than 200 samples, which is close to the designed initial data set for offline learning. Combined with the comparison of control performance in Fig. 3, where event-triggered online cooperative learning mechanism achieves smaller ultimate tracking error bound, the improved data efficiency is shown by choosing more informative data related to the control task using event-triggered online learning.

B. Manipulators with 2 Degrees of Freedoms

In this subsection, a multi-agent system consisting of $N=6$ manipulators with 2 degrees of freedoms (DoFs) is considered, where the communication topology is described as $\mathcal{G}=\{\mathcal{V},\mathcal{E}\}$ with $\mathcal{V}=\{1,\cdots,6\}$ and $\mathcal{E}=\{(1,2),(1,6),(2,3),(3,2),(4,3),(4,5),(5,6),(6,5)\}$. Each manipulator follows a second order dynamics in (1) with $\boldsymbol{g}(\boldsymbol{x}_i)=\boldsymbol{M}^{-1}(\boldsymbol{x}_{i,1})\boldsymbol{T}(\boldsymbol{x}_{i,1})$, $\boldsymbol{h}(\boldsymbol{x}_i)=\boldsymbol{M}^{-1}(\boldsymbol{x}_{i,1})\boldsymbol{f}_h(\boldsymbol{x}_i)$ and $n=2$, $p=q=2$, where $\boldsymbol{x}_{i,1},\boldsymbol{x}_{i,2}$ represent the absolute position and translational velocity of end effector for each manipulator i in inertial coordinate with the compact domain $\mathbb{X}=[-1.5,1.5]^4$. The mass matrix $\boldsymbol{M}(\boldsymbol{x}_{i,1})$, Jacobian matrix $\boldsymbol{T}(\boldsymbol{x}_{i,1})$ and generalized force $\boldsymbol{f}_h(\boldsymbol{x}_i)$ are known and calculated in appendix with corresponding manipulators' properties. The unknown function $\boldsymbol{f}(\cdot)$ represents the friction effect directly affected on the end effector and is written as

$$\boldsymbol{f}(\boldsymbol{x}_i)=\begin{bmatrix}5\sin(x_{i,1,1})+3\cos(x_{i,1,1})+x_{i,2,1}^2+6\\3\cos(x_{i,1,2})+5\cos(x_{i,1,2})+x_{i,2,2}^2+10\end{bmatrix},$$

where $x_{i,j,k}$ denotes the k-th entry of $\boldsymbol{x}_{i,j}$ for $\forall k=1,\cdots,p$. Moreover, only agents 1 and 4 have access to the leader, indicating $b_{11}=b_{44}=1$ and $b_{ii}=0$ for $\forall i\in\mathcal{V}\backslash\{1,4\}$. The prediction of $\boldsymbol{f}(\cdot)$ on each agent is obtained using Gaussian process regression with squared exponential kernel as $\kappa_i(\boldsymbol{x}_i,\boldsymbol{x}_i')=\exp(2\|\boldsymbol{x}_i-\boldsymbol{x}_i'\|^2)$ for $\forall i\in\mathcal{V}$ and POE as aggregation strategy. The training data pairs satisfy Assumption 3 with the standard deviation of the measurement noise as $\sigma_{o,i}=0.01$ for $\forall i\in\mathcal{V}$. The grid factor and probability for calculating the uniform prediction error bound in Lemma 2 are set as $\tau=10^{-8}$ and $\delta=0.1$, respectively. The control objective is designed such that the end effector of each manipulator tracks a leader trajectory defined as

$$\boldsymbol{x}_{l,1}(t)=[\sin(0.5t),\cos(0.5t)]^T,\boldsymbol{x}_{l,2}(t)=\dot{\boldsymbol{x}}_{l,1}(t),\boldsymbol{x}_{l,r}(t)=\dot{\boldsymbol{x}}_{l,2}(t)$$

with relative states for each manipulator i as

$$\boldsymbol{s}_{i,1}(t)=0.2[\cos(1.5t+i\pi/3),\sin(1.5t+i\pi/3)]^T,$$
$$\boldsymbol{s}_{i,2}(t)=\dot{\boldsymbol{s}}_{i,1}(t),\qquad \boldsymbol{s}_{i,r}(t)=\dot{\boldsymbol{s}}_{i,2}(t).$$

The control parameters in (5) are set as $c=10$, $\boldsymbol{\lambda}=[1,1]^T$. Moreover, choose $\boldsymbol{Q}_{\varepsilon}=\boldsymbol{I}_2$ inducing positive definite $\boldsymbol{Q}_z$ in (28). The simulation time is set from 0 to 40.

To demonstrate the effectiveness of the proposed event-triggered online learning strategies, the tracking error $\|\boldsymbol{\vartheta}(t)\|$ and the number of trigger events are compared among the learning strategies in Section V-A. Note that the number of initial data set for offline learning increases to 350 due to the expansion of the system dimension p. The comparison results are shown in the following subsections.

1) Performance with Event-triggered Online Learning: To show the effectiveness of different cooperative learning strategies, the tracking errors from the same initial states, i.e., $\boldsymbol{x}_1(0)=[0.8147,0.9058,0.1270,0.9134]^T$, $\boldsymbol{x}_2(0)=[0.6324,0.0975,0.2785,0.5469]^T$, $\boldsymbol{x}_3(0)=[0.9575,0.9649,0.1576,0.9706]^T$, $\boldsymbol{x}_4(0)=[0.9572,0.4854,0.8003,0.1419]^T$, $\boldsymbol{x}_5(0)=[0.4218,0.9157,0.7922,0.9595]^T$, $\boldsymbol{x}_6(0)=[0.6557,0.0357,0.8491,0.9340]^T$, are compared. The tracking error $\|\boldsymbol{\vartheta}_i(t)\|$ for each agent $i\in\mathcal{V}$ are shown in Fig. 5. It is obvious to see online learning performs better with lower $\|\boldsymbol{\vartheta}_i(t)\|$ for $\forall i\in\mathcal{V}$, while with offline learning the tracking error is much larger than with other methods especially due to the lack of sufficient data in the domain around references. Note that the tracking error does not tend to 0 due to the non-fully connection between the agents and leader inducing

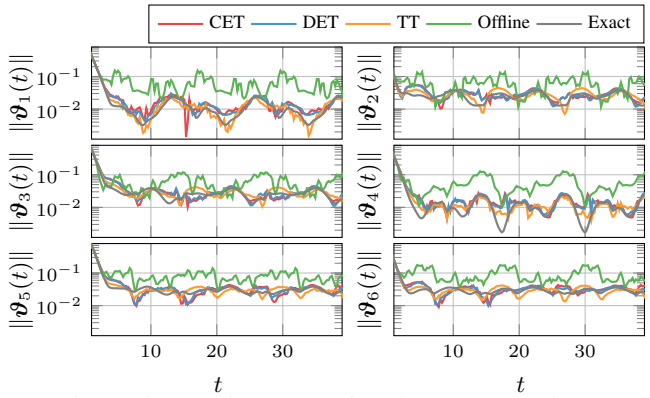


Figure 5: Tracking error of each agent over time.

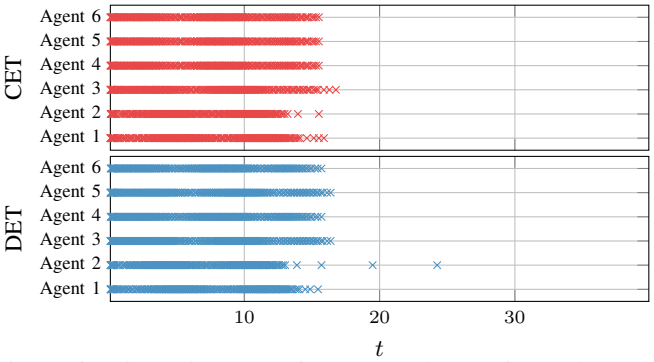


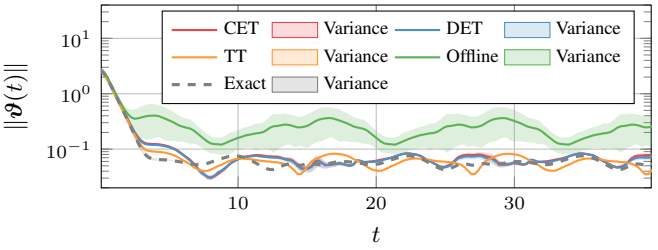
Figure 6: Trigger instances for CET and DET for each agent. Specifically, for CET 250, 238, 273, 263, 277 and 260 samples are collected for agent 1 to 6. With DET, each agent collects 274, 258, 303, 288, 303 and 286 data pairs.

non-zeros $\boldsymbol{\iota}$ in Theorem 1. However, the tracking performance of CET, DET and TT is similar, showing the efficiency of the event-triggered strategies.

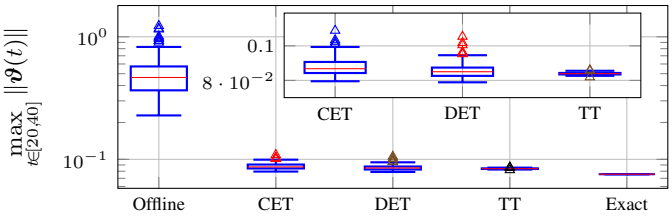
The time instances for each trigger in centralized and distributed event-triggered online learning are shown in Fig. 6 for each agent. It is seen that total number of data pairs in each agent is smaller than 350, which is the size of initial data set for offline learning. Moreover, no trigger occurs after $t = 25$, indicating sufficiently accurate GP prediction is achieved for the entire MAS and desired tracking performance.

2) Monte Carlo Test: The statistical property of the derived theorems are demonstrated through Monte Carlo Test, where the simulations are repeated for 100 times with random initial states $\boldsymbol{x}(0)$ sampled from uniform distribution in $[0, 1]^{nN_p}$, normally distributed measurement noise and random initial data set with $\boldsymbol{x}_i^{(\iota)} \in \mathbb{X}$, $\forall i \in \mathcal{V}$, $\forall \iota = 1, \cdots, 350$ for offline learning. The control performance for each learning algorithm is reflected by the ultimate tracking error as shown in Fig. 7. Specifically, the overall tracking error $\|\boldsymbol{\vartheta}(t)\|$ over time t is shown in Fig. 7a. Moreover, Fig. 7b indicates the ultimate tracking error by considering the maximal $\|\boldsymbol{\vartheta}(t)\|$ after $t = 20$, such that the initial behavior of the systems is excluded.

The online learning efficiency is reflected by the maximal size of eventual training data set on each agent shown in Fig. 8. It is easy to see that the maximal number of data pairs in the training data set for each agent with both event-triggered online learning mechanisms results is less than 350 samples, which is the designed size of initial data set for offline



(a) Mean and variance of overall tracking error with respect to time from 100 times Monte Carlo tests.



(b) Maximal tracking error in the steady state.

Figure 7: Tracking error from different learning strategy.

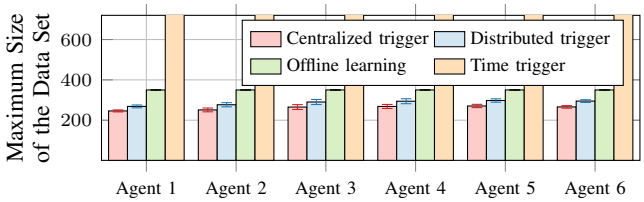


Figure 8: Maximal size of data set, which for offline learning is 350 from the size of initial data set and for TT is 2667 due to the fixed trigger interval 0.015. With CET, the maximal size of the data set for each agent 1 to 6 denotes 246 ± 5, 251 ± 9, 265 ± 12, 268 ± 10, 271 ± 8 and 266 ± 6, respectively. The maximal numbers of training samples collected through DET are 268 ± 7, 277 ± 10, 290 ± 13, 294 ± 12, 298 ± 8 and 295 ± 7 for agent 1 to 6, respectively.

learning. Moreover, in comparison to centralized methods, the distributed approach allows agents to collect only a slightly larger amount of data without the need for an extra computational center. Combining with the observation in Fig. 6, limited data is eventually added into individual data set for each agent with event-triggered strategies. In comparison, time-triggered method intends to collect data without termination, which results in an infinitely growing data set linear with operating time and high demand on the computational resources.

VI. Conclusions

In this paper, we consider a distributed control framework for leader-following time-varying formation control with cooperative online learning algorithm using GP regression. For high data efficiency with guaranteed control performance, two event-triggered mechanisms are proposed for online learning, namely the centralized and distributed version. Inspired from the centralized event-triggered learning design, the distributed trigger condition is evaluated only based on the local and neighboring information on each agent. Moreover, we show the exclusion of the Zeno behavior on each agent for both centralized and distributed event-trigger. Finally, the effectiveness of both proposed event-triggered cooperative online

learning strategies is demonstrated in the simulations, which show that each agent collects fewer data and triggers less GP model update compared with other approaches while achieving a guaranteed control performance.

APPENDIX

Proof of Proposition 1: The proof follows [28], where the sign of $\dot{V}$ for $\|\boldsymbol{z}\| > \chi^{-1}\bar{\vartheta}_c$ is investigated such that (39) is reformulated as $\bar{\rho} = \|\boldsymbol{z}\|$. Then, two cases divided by (39) is considered. In the case with $\rho(t) < \bar{\rho}(t)$ indicating no model update, the negativity of $\dot{V}$ is shown as $\dot{V} < -\underline{\lambda}(\boldsymbol{Q}_z)\|\boldsymbol{z}\|(\xi\rho - \xi\|\boldsymbol{\iota} + \hat{\boldsymbol{\eta}}_\delta(\boldsymbol{x})\|) \leq 0$ by considering (34) with (35). If $\rho(t) \geq \bar{\rho}(t)$, then the model update is activated, and with $\|\boldsymbol{z}\| > \chi^{-1}\bar{\vartheta}_c$ it has

$$\begin{aligned}\dot{V} &< -\underline{\lambda}(\boldsymbol{Q}_z)\|\boldsymbol{z}\|(\chi^{-1}\bar{\vartheta}_c - \xi\|\boldsymbol{\iota} + \tilde{\underline{\boldsymbol{\eta}}}_\delta(\boldsymbol{x})\|) \\ &\leq -\underline{\lambda}(\boldsymbol{Q}_z)\|\boldsymbol{z}\|\xi(\|\boldsymbol{\iota} + \underline{\hat{\boldsymbol{\eta}}}_\delta + \boldsymbol{\epsilon}\| - \|\boldsymbol{\iota} + \tilde{\underline{\boldsymbol{\eta}}}_\delta(\boldsymbol{x}) + \boldsymbol{\epsilon}\|) \leq 0,\end{aligned}$$

where the second inequality is derived by using the definition of $\bar{\vartheta}_c$ and the result from (40). Until here, the negativity of $\dot{V}$ when $\|\boldsymbol{z}\| > \chi^{-1}\bar{\vartheta}_c$ is proven, which concludes the proof by letting $\bar{z} = \chi^{-1}\bar{\vartheta}_c$ and using the result in Lemma 4. ∎

Proof of Property 1: (i) The property for agent $i \in \mathbb{T}$ is obvious, since the GP model update is activated on agent i such that $\tilde{\eta}_{\delta,i}(\boldsymbol{x}_i) = \hat{\eta}^+_{\delta,i}(\boldsymbol{x}_i) \leq \underline{\hat{\eta}}_{\delta,i} \leq \underline{\hat{\eta}}_{\delta,i} + \epsilon_i$ by considering Lemma 3 and positive ϵ_i. To prove the property for agent $i \in \mathbb{S} \cap \bar{\mathbb{T}}$, the trigger condition is investigated for $i \in \bar{\mathbb{T}}$ as $\rho_i \leq \bar{\rho}_i = (\underline{\hat{\eta}}_{\delta,i} + \iota_i + \epsilon_i)^2$ due to the definition of $\mathbb{S}$, which leads to $\tilde{\eta}_{\delta,i}(\boldsymbol{x}_i) = \hat{\eta}_{\delta,i}(\boldsymbol{x}_i) \leq \underline{\hat{\eta}}_{\delta,i} + \epsilon_i$ from the design of ρ_i in (43).

(ii) This property is directly derived by the definition of $\bar{\mathbb{S}}$, which is recalled as $\|\boldsymbol{z}_i\| > \chi^{-1}\bar{\vartheta}/\sqrt{N}$.

(iii) Considering the definition of $\bar{\mathbb{S}}$ and $\bar{\mathbb{T}}$, it has for agent $i \in \bar{\mathbb{S}} \cap \bar{\mathbb{T}}$ as $\rho_i \leq \xi^{-2}\|\boldsymbol{z}_i\|^2 - \chi^{-2}\bar{\vartheta}_d^2/N + \left(\underline{\hat{\eta}}_{\delta,i} + \iota_i + \epsilon_i\right)^2$, which is equivalent to Property 1 (3). ∎

REFERENCES

[1] J. Cui, Y. Liu, and A. Nallanathan, "Multi-agent Reinforcement Learning-based Resource Allocation for UAV Networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 729–743, 2019.

[2] Z. Yan, Z. Yang, A. Jiang, Y. Teng, X. Liu, and S. Wei, "Coordinated Control for Trajectory Tracking of Multiple UUVs with Input Saturation," in *OCEANS 2019-Marseille*. IEEE, 2019, pp. 1–5.

[3] X. Dai, H. Wu, S. Wang, J. Jiao, G. T. Nguyen, F. H. Fitzek, and S. Hirche, "Fast IMU-based Dual Estimation of Human Motion and Kinematic Parameters via Progressive In-Network Computing," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 8875–8882, 2023.

[4] Y. Yang, H. Modares, K. G. Vamvoudakis, and F. L. Lewis, "Cooperative Finitely Excited Learning for Dynamical Games," *IEEE Trans. Cybern.*, 2023.

[5] G. Liu, Y. Yang, X. Zhao, and C. K. Ahn, "Adaptive Fuzzy Practical Bipartite Synchronization for Multi-Agent Systems With Intermittent Feedback Under Multiple Unknown Control Directions," *IEEE Trans. Fuzzy Syst.*, 2024.

[6] C. K. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.

[7] Y. He and Y. Zhao, "Adaptive Robust Control of Uncertain Euler–Lagrange Systems Using Gaussian Processes," *IEEE Trans. Neural Networks Learn. Syst.*, 2022.

[8] S. Zhang, D.-H. Zhai, Y. Xiong, J. Lin, and Y. Xia, "Safety-Critical Control for Robotic Systems with Uncertain Model via Control Barrier Function," *Int. J. Robust Nonlinear Control*, vol. 33, no. 6, pp. 3661–3676, 2023.

[9] T.-Y. Huang, S. Zhang, X. Dai, A. Capone, V. Todorovski, S. Sosnowski, and S. Hirche, "Learning-based Prescribed-time Safety for Control of Unknown Systems with Control Barrier Functions," *IEEE Control Systems Letters*, 2024.

[10] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3250–3265, 2012.

[11] H. Liu, Y.-S. Ong, X. Shen, and J. Cai, "When Gaussian process Meets Big Data: A Review of Scalable GPs," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 31, no. 11, pp. 4405–4423, 2020.

[12] Z. Yang, S. Sosnowski, Q. Liu, J. Jiao, A. Lederer, and S. Hirche, "Distributed Learning Consensus Control for Unknown Nonlinear Multi-agent Systems based on Gaussian Processes," in *IEEE Conference on Decision and Control*. IEEE, 2021, pp. 4406–4411.

[13] Y. Cao and D. J. Fleet, "Generalized Product of Experts for Automatic and Principled Fusion of Gaussian Process Predictions," *arXiv preprint arXiv:1410.7827*, 2014.

[14] Z. Yang, X. Dai, A. Dubey, S. Hirche, and G. Hattab, "Whom to Trust? Elective Learning for Distributed Gaussian Process Regression," in *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*, 2024, pp. 2020–2028.

[15] Z. Yang, S. Dong, A. Lederer, X. Dai, S. Chen, S. Sosnowski, G. Hattab, and S. Hirche, "Cooperative Learning with Gaussian Processes for Euler-Lagrange Systems Tracking Control under Switching Topologies," *arXiv preprint arXiv:2402.03048*, 2024.

[16] A. Lederer, Z. Yang, J. Jiao, and S. Hirche, "Cooperative Control of Uncertain Multi-Agent Systems via Distributed Gaussian Processes," *IEEE Trans. Autom. Control*, 2022.

[17] T. N. Hoang, Q. M. Hoang, K. H. Low, and J. How, "Collective Online Learning of Gaussian Processes in Massive Multi-agent Systems," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 7850–7857.

[18] W. Wang, X. Yue, B. Haaland, and C. Jeff Wu, "Gaussian Processes with Input Location Error and Applications to the Composite Parts Assembly Process," *SIAM/ASA J. Uncertainty Quantif.*, vol. 10, no. 2, pp. 619–650, 2022.

[19] J. Umlauft and S. Hirche, "Feedback Linearization based on Gaussian Processes with Event-triggered Online Learning," *IEEE Trans. Autom. Control* , vol. 65, no. 10, pp. 4154–4169, 2019.

[20] T. Beckers, S. Hirche, and L. Colombo, "Online Learning-based Formation Control of Multi-agent Systems with Gaussian Processes," in *IEEE Conference on Decision and Control*, 2021, pp. 2197–2202.

[21] T. Beckers, G. J. Pappas, and L. J. Colombo, "Learning Rigidity-based Flocking Control using Gaussian Processes with Probabilistic Stability Guarantees," in *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 7254–7259.

[22] X. Dai, Z. Yang, M. Xu, S. Zhang, F. Liu, G. Hattab, and S. Hirche, "Decentralized Event-triggered Online Learning for Safe Consensus Control of Multi-agent Systems with Gaussian Process Regression," *European Journal of Control*, p. 101058, 2024.

[23] F. Solowjow, D. Baumann, J. Garcke, and S. Trimpe, "Event-Triggered Learning for Resource-Efficient Networked Control," in *Proceedings of the American Control Conference*, 2018, pp. 6506–6512.

[24] F. Castañeda, J. J. Choi, W. Jung, B. Zhang, C. J. Tomlin, and K. Sreenath, "Probabilistic Safe Online Learning with Control Barrier Functions," *arXiv preprint arXiv:2208.10733*, 2022.

[25] J. Zhang, H. Zhang, Z. Ming, and Y. Mu, "Adaptive Event-triggered Time-varying Output Bipartite Formation Containment of Multiagent Systems under Directed Graphs," *IEEE Trans. Neural Networks Learn. Syst.*, 2022.

[26] H. Liang, G. Liu, H. Zhang, and T. Huang, "Neural-network-based Event-triggered Adaptive Control of Nonaffine Nonlinear Multiagent Systems with Dynamic Uncertainties," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 32, no. 5, pp. 2239–2250, 2020.

[27] F. Gao, W. Chen, Z. Li, J. Li, and B. Xu, "Neural Network-based Distributed Cooperative Learning Control for Multiagent Systems via Event-triggered Communication," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 31, no. 2, pp. 407–419, 2019.

[28] X. Dai, A. Lederer, Z. Yang, and S. Hirche, "Can Learning Deteriorate Control? Analyzing Computational Delays in Gaussian Process-Based Event-Triggered Online Learning," in *Learning for Dynamics and Control Conference*. PMLR, 2023, pp. 445–457.

[29] J. Jiao, A. Capone, and S. Hirche, "Backstepping Tracking Control using Gaussian Processes with Event-triggered Online Learning," *IEEE Control Syst. Lett.*, vol. 6, pp. 3176–3181, 2022.

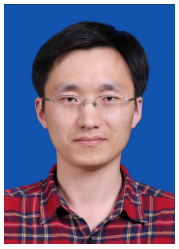
[30] H. K. Khalil, *Nonlinear Systems*. Prentice-Hall, 2002.

[31] W. Ren and R. Beard, "Consensus Seeking in Multiagent Systems under Dynamically Changing Interaction Topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, 2005.
[32] H. Zhang and F. L. Lewis, "Adaptive Cooperative Tracking Control of Higher-order Nonlinear Systems with Unknown Dynamics," *Automatica*, vol. 48, no. 7, pp. 1432–1439, 2012.
[33] Z. Zuo, B. Tian, M. Defoort, and Z. Ding, "Fixed-time Consensus Tracking for Multiagent Systems with High-order Integrator Dynamics," *IEEE Trans. Autom. Control*, vol. 63, no. 2, pp. 563–570, 2017.
[34] L. Gao, X. Dai, M. Kleeberger, and J. Fottner, "Quasi-static Optimal Control Strategy of Lattice Boom Crane based on Large-Scale Flexible Non-linear Dynamics," in *Simulation and Modeling Methodologies, Technologies and Applications: International Online Conference (SIMULTECH 2021)*. Springer, 2023, pp. 153–177.
[35] S. Zhang, D.-H. Zhai, J. Lin, Y. Xiong, Y. Xia, and M. Wei, "ESO-Based Safety-Critical Control for Robotic Systems With Unmeasured Velocity and Input Delay," *IEEE Trans. Ind. Electron.*, 2024.
[36] C. Kaddissi, J.-P. Kenne, and M. Saad, "Identification and Real-time Control of an Electrohydraulic Servo System based on Nonlinear Backstepping," *IEEE/ASME Trans. Mechatron.*, vol. 12, no. 1, pp. 12–22, 2007.
[37] A. Luo, Q. Zhou, H. Ma, and H. Li, "Observer-based Consensus Control for MASs with Prescribed Constraints via Reinforcement Learning Algorithm," *IEEE Trans. Neural Networks Learn. Syst.*, 2023.
[38] Q. Zhu, B. Niu, D. Wang, S. Li, X. Wang, and J. Kong, "Cooperative ETM-based Adaptive Neural Network Tracking Control for Nonlinear Pure-feedback Mass: A Special-shaped Laplacian Matrix Method," *IEEE Trans. Neural Networks Learn. Syst.*, 2022.
[39] W. Bai, P. X. Liu, and H. Wang, "Neural-network-based adaptive fixed-time control for nonlinear multiagent non-affine systems," *IEEE Trans. Neural Networks Learn. Syst.*, 2022.
[40] A. Lederer, J. Umlauft, and S. Hirche, "Uniform Error Bounds for Gaussian Process Regression with Application to Safe Control," *Advances in Neural Information Processing Systems*, vol. 32, 2019.
[41] S. Zhang, D.-H. Zhai, X. Dai, T. yuan Huang, Y. Xia, and S. Hirche, "Learning-based Parameterized Barrier Function for Safety-Critical Control of Unknown Systems," 2024.
[42] A. Lederer, A. J. O. Conejo, K. A. Maier, W. Xiao, J. Umlauft, and S. Hirche, "Gaussian Process-based Real-time Learning for Safety Critical Applications," in *International Conference on Machine Learning*. PMLR, 2021, pp. 6055–6064.
[43] V. Tresp, "Mixtures of Gaussian processes," *Advances in Neural Information Processing Systems*, vol. 13, 2000.
[44] Y. Fan, L. Liu, G. Feng, and Y. Wang, "Self-triggered Consensus for Multi-agent Systems with Zeno-free Triggers," *IEEE Trans. Autom. Control*, vol. 60, no. 10, pp. 2779–2784.
[45] X. Dai, Z. Yang, and S. Hirche, "Cooperative Online Learning for Multi-Agent System Control via Gaussian Processes with Event-Triggered Mechanism: Extended Version," *arXiv preprint arXiv:2304.05138*, 2023.

Xiaobing Dai received the B.Sc. mechanical engineering from the Tongji University, Shanghai, China, in 2018 with direction in mechatronics, building environment and civil engineering. He received double M.Sc degrees in Mechanical Engineering, Mechatronics and Robotics from the Technical University of Munich, Munich, Germany, in 2021. Since February 2022, he is a PhD student at the Chair of Information-oriented Control, TUM School of Computation, Information and Technology at the Technical University of Munich, Germany. His current research interests include efficient machine learning, networked control systems, safe learning-based control.

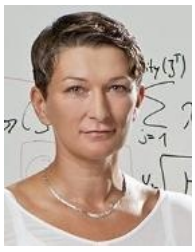
Zewen Yang received the M.S. degree in control engineering from Northeast Forest University, in 2017. He pursued a Ph.D. in control science and engineering at College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin, China, from 2017 to 2019 and joined the Chair of Information-oriented Control, School of Computation, Information and Technology, the Technical University of Munich, Munich, Germany, in machine learning and data-driven control until 2023. His current research interests include multi-agent systems, cooperative learning, control theory, and general robotics.

Sihua Zhang received the M.Eng. degree in control engineering from the Beijing Institute of Technology, Beijing, China, in 2021. She is currently pursuing the doctor's degree in control engineering with the School of Automation, Beijing Institute of Technology, Beijing, China. From 2023 to 2025, she is additionally working as research assistant at the Chair of Information-oriented Control, TUM School of Computation, Information and Technology at the Technical University of Munich, Germany. Her current research interests include safety-critical robotic control, control barrier functions, machine learning and optimal control.



Di-Hua Zhai received the B.Eng. degree in automation from Anhui University, Hefei, China, in 2010, the M.Eng. degree in control science and engineering from the University of Science and Technology of China, Hefei, in 2013, and the Dr.Eng. degree in control science and engineering from the Beijing Institute of Technology, Beijing, China, in 2017. Since 2017, he has been with the School of Automation, Beijing Institute of Technology, where he is currently an Associate Professor. His research interests include teleoperation, intelligent robot, human robot collaboration, switched control, optimal control, constrained control, and networked control.

Yuanqing Xia (Fellow, IEEE) received his M.S. degree in Fundamental Mathematics from Anhui University, China, in 1998 and his Ph.D. degree in Control Theory and Control Engineering from Beijing University of Aeronautics and Astronautics, Beijing, China, in 2001. During January 2002-November 2003, he was a Postdoctoral Research Associate with the Institute of Systems Science, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing, China. From November 2003 to February 2004, he was with the National University of Singapore as a Research Fellow, where he worked on variable structure control. From February 2004 to February 2006, he was with the University of Glamorgan, Pontypridd, U.K., as a Research Fellow. From February 2007 to June 2008, he was a Guest Professor with Innsbruck Medical University, Innsbruck, Austria. Since 2004, he has been with the Department of Automatic Control, Beijing Institute of Technology, Beijing, first as an Associate Professor, then, since 2008, as a Professor. His current research interests are in the fields of cloud control systems, networked control systems, robust control and signal processing, active disturbance rejection control, unmanned system control, and flight control.



Sandra Hirche (M'03–SM'11–F'20) received the Dipl.-Ing degree in aeronautical engineering from the Technical University of Berlin, Germany, in 2002, and the Dr. Ing. degree in electrical engineering from the Technical University of Munich (TUM), Germany, in 2005. From 2005 to 2007, she was awarded a Post-doctoral scholarship from the Japanese Society for the Promotion of Science at the Fujita Laboratory, Tokyo Institute of Technology, Japan. From 2008 to 2012, she was an Associate Professor with TUM. Since 2013, she has served as TUM Liesel Beckmann Distinguished Professor and has been with the Chair of Information-Oriented Control, TUM. She has authored or coauthored more than 150 papers in international journals. Her main research interests include cooperative, distributed, and networked control with applications in human–machine interaction, multirobot systems, and general robotics.

Dr. Hirche has served on the editorial boards of the IEEE Transactions on Control of Network Systems, the IEEE Transactions on Control Systems Technology. She has received multiple awards such as the Rohde & Schwarz Award for her Ph.D. thesis, the IFAC World Congress Best Poster Award in 2005, and the Best Paper Awards from IEEE Worldhaptics and the IFAC Conference of Manoeuvring and Control of Marine Craft in 2009.