Decentralized Event-Triggered Online Learning for Safe Consensus of Multi-Agent Systems with Gaussian Process Regression

Xiaobing Dai[*1], Zewen Yang[*†2], Mengtian Xu[1], Sihua Zhang[3],
Fangzhou Liu[4], Georges Hattab[2,5] and Sandra Hirche[1]

Abstract— Consensus control in multi-agent systems has received significant attention and practical implementation across various domains. However, managing consensus control under unknown dynamics remains a significant challenge for control design due to system uncertainties and environmental disturbances. This paper presents a novel learning-based distributed control law, augmented by an auxiliary dynamics. Gaussian processes are harnessed to compensate for the unknown components of the multi-agent system. For continuous enhancement in predictive performance of Gaussian process model, a data-efficient online learning strategy with a decentralized event-triggered mechanism is proposed. Furthermore, the control performance of the proposed approach is ensured via the Lyapunov theory, based on a probabilistic guarantee for prediction error bounds. To demonstrate the efficacy of the proposed learning-based controller, a comparative analysis is conducted, contrasting it with both conventional distributed control laws and offline learning methodologies.

I. INTRODUCTION

In recent decades, multi-agent systems (MASs) have garnered considerable interest due to their versatile applications, encompassing domains such as underwater vehicles, electrical machines and multi-zone buildings [1]–[4]. Given the inherent challenges associated with uncertainties arising from environmental disturbances and inaccurate system dynamics, learning-based control approaches have emerged as a promising avenue for mitigating these issues. These approaches harness machine learning methodologies to estimate unknown components in the system dynamics, drawing from observed data. Subsequently, these data-driven models are seamlessly integrated into the control systems to enhance their adaptability.

The substantial progress witnessed in the field of deep learning, particularly its capacity to model intricate mappings from data, has unequivocally propelled the adoption of neural networks (NNs) for regression tasks. It is evident that the widespread success of NNs in diverse applications has instigated their integration into the domain of consensus control in MASs [5]. For instance, an adaptive protocol that incorporates data-driven models acquired through NNs has been proposed in [6], [7]. Furthermore, research efforts have delved into security-related issues, including scenarios involving attacks and actuator faults within compromised networks [8], [9]. However, as the number of hidden layers and neurons in deep learning models increases, the adjustment of model weights becomes a time-consuming process. This poses a significant practical challenge for real-time implementation in control systems, primarily because the time delay associated with the learning process adversely affects control performance [10], [11]. Additionally, the integration of NNs does not inherently provide precise prediction guarantees, a limitation that hinders their application in safety-critical domains.

In contrast, Gaussian process (GP) regression [12] has emerged as an alternative approach that carries certain advantages in the context of MAS control. GP provides a probabilistic framework that facilitates principled uncertainty quantification, offering a valuable tool for modeling uncertainties within the MAS and enabling safer control strategies [13]. GP regression is first introduced for cooperative learning-based control in MASs in [14], which conceives the novel idea of treating each agent as an individual GP model. However, the aggregation strategy requires the computation of the predictions for neighbor agents, resulting in a larger demand for local computational resources. To reduce the computational burden, subsequently, a dynamic average consensus algorithm [15] is integrated to enhance the posterior mean and variance of GPs with only prediction on the local states for each agent [16]. But the aforementioned works overlook the imperative consideration of online learning with MAS control, leading to a deficiency in the ability to update GP models and a restricted efficacy in prediction. To address this issue, an online learning approach by adding new training samples was proposed for MAS flocking control in [17]. However, it only considers the data collection in a fixed frequency with the deletion strategy for data storage management and lacks the consideration of the different importance of the data. To tackle this problem, an event-triggered learning approach for formation control of MASs is presented in [18] to choose only valuable data for GP-based control. Nevertheless, each agent requires access to global

*Equal contribution. †Corresponding author.

[1]Chair of Information-oriented Control, TUM School of Computation, Information and Technology, Technical University of Munich, 80333 Munich, Germany. `xiaobing.dai; mengtian.xu; hirche@tum.de` [2]Center for Artificial Intelligence in Public Health Research (ZKI-PH), Robert Koch Institute, Berlin, Germany. `yangz; hattabg@rki.de` [3]School of Automation, Beijing Institute of Technology, Beijing, P.R.China. `sihua.zhang@bit.edu.cn` [4]National Key Laboratory of Modeling and Simulation for Complex Systems, Harbin Institute of Technology, 150001, P.R.China. `fangzhou.liu@hit.edu.cn` [5]Department of Mathematics and Computer Science, Freie Universität Berlin, Berlin, Germany.

information such as communication topologies, which may not be feasible in certain circumstances.

In this paper, we consider the average consensus for a multi-agent system with unknown dynamics. A novel distributed learning-based consensus controller featuring an auxiliary dynamics is proposed, which isolates the influence of the dynamics uncertainties from the estimation of the average states. Moreover, Gaussian process regression is employed to compensate for the unknown dynamics. For better control performance by using the GP-based controller, a decentralized event-triggered online learning strategy is introduced for control performance improvement with efficient data storage management. Furthermore, a stability analysis of the MAS is included with the probabilistic prediction guarantees for average consensus error, further enhancing its operational effectiveness for safe learning-based control tasks. To underscore the superiority of the proposed approach, numerical simulations are carried out demonstrating its effectiveness through a comparative analysis against conventional control and learning strategies.

The remaining content of this paper is structured as follows. Section II presents the problem setting and proposes the distributed consensus control law. In Section III, the development of the decentralized event-triggered strategy for online learning is provided with control performance guarantee. Section IV encompasses the execution of numerical simulations, followed by the conclusion of this paper in Section V.

II. PROBLEM SETTING AND PRELIMINARIES

In this paper, a MAS with $N \in \mathbb{N}_+$ homogeneous agents is considered, whose dynamics is expressed by

$$\dot{x}_i = h(x_i) + f(x_i) + g(x_i)u_i, \ \forall i = 1, \cdots, N, \quad (1)$$

where $x_i \in \mathbb{X} \subset \mathbb{R}$ with a compact domain $\mathbb{X}$ and $u_i \in \mathbb{R}$ denote the system state and control input, respectively. For notational simplicity, we focus on scalar agent dynamics, whose results can be directly extended to multi-dimensional cases by using multi-input-output machine learning tools and the Kronecker operator. The functions $h : \mathbb{X} \to \mathbb{R}$ and $g : \mathbb{X} \to \mathbb{R}$ describe the known parts of the individual dynamics of the agents, while the function $f : \mathbb{X} \to \mathbb{R}$ is regarded as unknown components such as unmodeled dynamics and environmental uncertainties. To ensure the controllability for each agent, the following assumption on $g(\cdot)$ is required.

Assumption 1: For any $x_i \in \mathbb{X}$, it has $g(x_i) \neq 0$.
The non-singularity of $g(\cdot)$ is a common prerequisite for the controller non-autonomous system, such that the control input u_i is effective at any state. While this assumption induces restrictions on the system class, the focus of this paper is on the design of event-triggered learning-based controllers. The extension to other system classes is left in the future work.

The agents communicate through a network, whose topology is presented by a time-invariant undirected connected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{1, \cdots, N\}$ is the vertex set representing the indices of agents and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denotes the edge set. The topology of $\mathcal{G}$ is characterized by an adjacency matrix $\mathcal{A} = [a_{ij}]_{i,j \in \mathcal{V}} \in \mathbb{R}^{N \times N}$. The

entry $a_{ij} = a_{ji} = 1$, when there exists a communication channel such that the agent i can send and receive the information from agent j, i.e., $(i, j) \in \mathcal{E}$, otherwise $a_{ij} = a_{ji} = 0, \forall i, j \in \mathcal{V}$. The set $\mathcal{N}_i$ of agent i contains all the neighboring agents $j \in \mathcal{V}$ indicating $a_{ij} = 1$. The Laplacian matrix of $\mathcal{G}$ is defined as $\mathcal{L} = [l_{ij}]_{i,j \in \mathcal{V}} \in \mathbb{R}^{N \times N}$ with $l_{ii} = \sum_{j=1}^{N} a_{ij}$ and $l_{ij} = -a_{ij}$ for $j \neq i$.

The control task is to achieve average consensus, i.e., the agent's state x_i tends to reach $\bar{x}^* = N^{-1} \mathbf{1}_N^T \boldsymbol{x}(0)$ for $i = 1, \cdots, N$, where $\boldsymbol{x} = [x_1, \cdots, x_N]^T$ denotes the concatenated state. However, due to the unknown function $f(\cdot)$, the asymptotic stability of the MAS is not expected. Therefore, in this paper, we investigate ϵ-average consensus instead, whose definition is given as follows.

Definition 1: If there exists a well-defined constant $\epsilon \in \mathbb{R}_{0,+}$, such that $\lim_{t \to \infty} \|\boldsymbol{x}(t) - \mathbf{1}_N \bar{x}^*\| \leq \epsilon$. Then, the multi-agent system achieves ϵ-average consensus.

To achieve such a control task, we first consider a conventional controller using consensus law r_i from [19] as

$$u_i = -g^{-1}(x_i) \left(h(x_i) + \hat{f}_i(x_i) + cr_i \right), \quad (2a)$$

$$r_i = \sum_{j \in \mathcal{N}_i} (x_i - x_j), \qquad \forall i \in \mathcal{V}, \quad (2b)$$

where $c \in \mathbb{R}_+$ is the control gain and $\hat{f}_i(x_i)$ is the prediction of unknown $f(x_i)$ from agent i, which is illustrated in Section III. While the controlled system achieves consensus by using (2) under dynamics uncertainties from unknown $f_i(\cdot)$ proven in [20], the state of each agent is not guaranteed in the compact domain $\mathbb{X}$ and not close to $\bar{x}^*$. Without rigorous proof, an example is shown in Appendix reflecting the potential instability of the controlled system with (2). To address this problem, a novel distributed control law is proposed integrating an auxiliary dynamics as follows

$$u_i = -g^{-1}(x_i) \left(h(x_i) + \hat{f}_i(x_i) + cr_i - \dot{\bar{x}}_i \right), \quad (3a)$$

$$r_i = \sum_{j \in \mathcal{N}_i} (\tilde{x}_i - \tilde{x}_j) + \tilde{x}_i, \quad (3b)$$

where the error between the current state x_i and auxiliary state $\bar{x}_i$ is denoted by $\tilde{x}_i = x_i - \bar{x}_i$. The auxiliary state $\bar{x}_i$ of the agent i is obtained by solving the following dynamics as

$$\dot{\bar{x}}_i = -\bar{c} \sum_{j \in \mathcal{N}_i} (\bar{x}_i - \bar{x}_j), \ \forall i \in \mathcal{V}, \quad (4)$$

with a gain $\bar{c} \in \mathbb{R}_+$ and the initial condition $\bar{x}_i(0) = x_i(0)$.

The estimation function $\hat{f}_i(\cdot)$ of the unknown function $f(\cdot)$ in (3) associated with the agent i is obtained from data-driven method using the individual data set $\mathbb{D}_i = \{x_i^{(\iota)}, y_i^{(\iota)}\}_{\iota \in \mathbb{N}}$ of each agent $i \in \mathcal{V}$, which satisfies the following assumption.

Assumption 2: The data pair $\{x_i^{(\iota)}, y_i^{(\iota)}\}$ is available on each agent $i \in \mathcal{V}$ at any time $t_i^{(\iota)} \in \mathbb{R}_{0,+}$ with $\iota \in \mathbb{N}$, where the measured output denotes $y_i^{(\iota)} = \dot{x}_i(t_i^{(\iota)}) - h(x_i(t_i^{(\iota)})) - g(x_i(t_i^{(\iota)}))u_i(t_i^{(\iota)}) + w_i^{(\iota)}$. The measurement noise $w_i^{(\iota)} \in \mathbb{R}$ follows a zero-mean, independent and identical Gaussian distribution with variance σ_n^2 and $\sigma_n \in \mathbb{R}_+$.
Assumption 2 indicates each agent can collect the data individually. Additionally, this assumption accommodates the

Gaussian noise in $y_i^{(\iota)}$, which often stems from the numerical approximation of $\dot{x}_i(t_i^{(\iota)})$, such as the utilization of finite difference methods. The relaxation of the noise distribution exists as discussed in works such as [21], [22]. Moreover, Assumption 2 provides the possibility for the continuous enhancement of $\hat{f}_i(\cdot)$ through ongoing online data collection.

In order to collect the necessary data pairs individually on each agent with high data efficiency, we introduce a decentralized event-triggered online learning strategy. This strategy ensures that the pair $\{x_i^{(\iota)}, y_i^{(\iota)}\}$ will be added into the training dataset $\mathbb{D}_i$ of the agent i at time $t_i^{(\iota)}$ in accordance with the following condition

$$t_i^{(\iota)} = \min\{t_i : t_i > t_i^{(\iota-1)} \ \wedge \ \rho(x_i(t_i), \bar{x}_i(t_i)) > 0\}, \quad (5)$$

where $\rho(\cdot, \cdot) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ indicates the universal trigger function for all agents by evaluated by using individual state information including the auxiliary state $\bar{x}_i$. The detailed expression of $\rho(\cdot)$ will be presented in Section III.

III. DECENTRALIZED EVENT-TRIGGERED ONLINE LEARNING WITH GAUSSIAN PROCESSES

A. Gaussian Process Regression

To infer the unknown $f(\cdot)$, Gaussian process regression is adopted, which induces a distribution of $f(\cdot)$ by the prior mean $m(\cdot) : \mathbb{X} \to \mathbb{R}$ and kernel function $\kappa(\cdot, \cdot) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_{0,+}$, i.e., $f(\cdot) \sim \mathcal{GP}(m(\cdot), \kappa(\cdot, \cdot))$. The prior mean $m(\cdot)$ reflects the known part of $f(\cdot)$, which for the dynamics in (1) is encoded in $h(\cdot)$. Therefore, $m(\cdot) = 0$ is set. The kernel function $\kappa(\cdot)$ indicates the covariance between two sample points and satisfies the following assumption.

Assumption 3: The continuous function $f(\cdot)$ is a sample from a Gaussian process $\mathcal{GP}(0, \kappa(\cdot, \cdot))$ with a Lipschitz kernel $\kappa(|x - x'|) = \kappa(x, x')$ w.r.t $|x - x'|$.

From this assumption, the prior distribution of the unknown function $f(\cdot)$ is defined. Moreover, the assumption on the kernel $\kappa(\cdot, \cdot)$ can be achieved by choosing any continuous function reflecting the continuity of $f(\cdot)$, and then its Lipschitz continuity is derived by considering the compact input domain $\mathbb{X}$. And the commonly used kernels, such as square exponential kernel, satisfy the Lipschitz condition. Therefore, this assumption is practically not restrictive.

Consider an individual data set $\mathbb{D}_i$ on agent $i \in \mathcal{V}$ containing $M_i \in \mathbb{N}$ training samples and satisfying Assumption 2, the posterior mean $\mu_i(\cdot)$ and variance $\sigma_i(\cdot)$ are obtained by using the Gaussian process model with Assumption 3 as

$$\mu_i(x_i) = \boldsymbol{k}_{X,i}^T (\boldsymbol{K}_i + \sigma_n^2 \boldsymbol{I}_M)^{-1} \boldsymbol{y},$$
$$\sigma_i^2(x_i) = \kappa(0) - \boldsymbol{k}_{X,i}^T (\boldsymbol{K}_i + \sigma_n^2 \boldsymbol{I}_M)^{-1} \boldsymbol{k}_{X,i},$$

where $\boldsymbol{k}_{X,i} = [\kappa(x_i, x_i^{(1)}), \cdots, \kappa(x_i, x_i^{(M)})]^T$, $\boldsymbol{K}_i = [\kappa(x_i^{(p)}, x_i^{(q)})]_{p,q=1,\cdots,M}$ and $\boldsymbol{y} = [y_i^{(i)}, \cdots, y_i^{(M)}]^T$. While the obtained posterior mean $\mu_i(\cdot)$ is adopted as the compensation of the unknown function $f(\cdot)$ in (3), i.e., $\hat{f}_i(\cdot) = \mu_i(\cdot)$, the posterior variance $\sigma_i(\cdot)$ is used to quantify the prediction error shown as follows.

Lemma 1 ([23]): Consider the Gaussian process model on agent i for the inference of an unknown function $f(\cdot)$

satisfying Assumption 3, which is based on a data set under Assumption 2. Choose $\delta \in (0, 1)\mathbb{R}$ and $\tau_i \in \mathbb{R}_+$ such that $\gamma_i \leq \min_{x_i \in \mathbb{X}} \sqrt{\beta} \sigma_i(x_i)$ with

$$\beta = 2\log((2\delta\tau_i)^{-1}(\max_{x \in \mathbb{X}} x - \min_{x \in \mathbb{X}} x) + \delta^{-1}),$$
$$\gamma_i = (L_f + L_{\mu,i} + \sqrt{\beta}L_\sigma)\tau_i,$$

where L_f, $L_{\mu,i}$, L_σ are the Lipschitz constants for $f(\cdot)$, $\mu_i(\cdot)$, $\sigma_i(\cdot)$, respectively. Then, the prediction error is uniformly bounded by

$$|f(x_i) - \mu_i(x_i)| \leq \eta_i(x_i) = 2\sqrt{\beta}\sigma_i(x_i), \ \forall x_i \in \mathbb{X},$$

with probability of at least $1 - \delta$.

Although with conservatism, Lemma 1 provides a numerical way to probabilistically bound the prediction error, which is commonly used for the control in safety-critical scenarios. Note that Lemma 1 requires the Lipschitz continuity of the unknown function $f(\cdot)$ represented by L_f, which is easily shown by considering the continuity of $f(\cdot)$ in Assumption 3 and the compact domain $\mathbb{X}$. The approximation of L_f can be obtained by using the data-driven methods as in [23]. Moreover, the Lipschitz constants $L_{\mu,i}$ and L_σ exist due to the Lipschitz kernel function $\kappa(\cdot)$ in Assumption 3, whose explicit computation follows [24].

B. Decentralized Event-Triggered Learning

Since the performance of a learning based controller incorporated with the data-driven model $\hat{f}(\cdot)$ in (3) is contingent upon the precision of its predictions [23], it is common to employ online learning techniques to enhance prediction accuracy for better control performance. To facilitate effective and efficient online learning, a decentralized event-trigger mechanism is designed to improve the efficiency of the learning process. This mechanism induces a careful selection of the training data relying on a trigger function denoted as $\rho(\cdot)$ in (5). The specific choice of this trigger function is a vital aspect of our approach, aimed at achieving data efficiency and improving the prediction performance of the system during real-time operation.

$$\rho(x_i, \bar{x}_i) = \eta_i(x_i) - \max\{c|x_i - \bar{x}_i| - \sqrt{N-1}\underline{\eta}, \underline{\eta}\}, \quad (6)$$

where $\underline{\eta} = 2\sqrt{\beta}\sigma_n$. The trigger function $\rho(\cdot)$ is composed of three parts, namely the current local prediction error $\eta_i(x_i)$ defined in Lemma 1, the constant term $\sqrt{N-1}\underline{\eta}$ and the error between actual state x_i and auxiliary state $\bar{x}_i$, i.e., $\tilde{x}_i$. Note that the time stamp t_i in the trigger function $\rho(\cdot)$ in (5) is dropped in (6) for notational simplicity.

Remark 1: While based on the distributed control as in (3), the operations for GP models on each agent, including prediction and online learning, are only based on the local information, i.e., x_i and $\bar{x}_i$. Therefore, the proposed event-trigger is decentralized similarly as in [25]. Moreover, the distributed control law (3) with auxiliary dynamics (4) and the decentralized event-trigger (5) only require the exchange of system state x_i and $\bar{x}_i$ among agents, which reduces the communication burden in network compared with [26]. Furthermore, it requires no computation for the predictions

of the neighbors, decreasing the local computational demand on each agent compared to [27].

Remark 2: The constant $\sqrt{N-1}\underline{\eta}$ is introduced, due to the lack of information about actions of online learning on other agents. Moreover, by considering its non-negativity because of $N \geq 1$ and $\sigma_n > 0$ as in Assumption 2, the event-trigger becomes more conservative than single agent case in [10], i.e., inducing more trigger events. Moreover, the conservatism grows with the number of agents N, which is intuitive due to more unknown actions on more agents. In particular for a single agent system as in [28], the trigger function $\rho(\cdot)$ in (6) is degenerated to $\rho(x_i, \bar{x}_i) = \eta_i(x_i) - \max\{c|\tilde{x}_i|, \underline{\eta}\}$. The degenerated form indicates the conditions $\rho(x_i, \bar{x}_i) > 0$ and $\{\eta_i(x_i) > c|\tilde{x}_i| \vee \eta_i(x_i) > \underline{\eta}\}$ are equivalent and identical as in [29], whose effects on the learning-based control are proven with ultimately bounded error.

When the trigger condition (5) is satisfied on agent i, the GP model and the prediction $\mu_i(x_i)$ on agent i will be updated by adding new training data. We denote the updated posterior mean, variance, and the corresponding prediction error by $\mu_i^+(x_i)$, $\sigma_i^+(x_i)$, and $\eta_i^+(x_i)$, respectively, using the updated data set satisfying

$$|f(x_i) - \mu_i^+(x_i)| \leq \eta_i^+(x_i) \leq \underline{\eta}, \ \forall x_i \in \mathbb{X}, \qquad (7)$$

with a probability of at least $1 - \delta$, which is proven by considering the 1-point upper bound $\sigma_{i,1\mathrm{pt}}^+$ of variance [30] as

$$(\sigma_{i,1\mathrm{pt}}^+)^2 = \kappa(0) - \frac{\kappa(0)^2}{\kappa(0) + \sigma_n^2} = \frac{\kappa(0)\sigma_n^2}{\kappa(0) + \sigma_n^2} \leq \sigma_n^2 \qquad (8)$$

and the fact that more data results in smaller posterior variance [10], i.e., $\sigma_i^+(x_i) \leq \sigma_{i,1\mathrm{pt}}^+ \leq \sigma_n$. Combining with the definition of the prediction error bound in Lemma 1, the upper bound of $\eta_i^+(x_i)$ is derived as in (7). For notational simplicity, define $\hat{\mu}_i(x_i) = \mu_i(x_i)$ and $\hat{\eta}_i(x_i) = \eta_i(x_i)$ for all agent i with $\rho(x_i, \bar{x}_i) \leq 0$, as well as $\hat{\mu}_i(x_i) = \mu_i^+(x_i)$ and $\hat{\eta}_i(x_i) = \eta_i^+(x_i)$ for all agent i with $\rho(x_i, \bar{x}_i) > 0$. Note that only the prediction error bound without model update, i.e., $\mu_i(x_i)$, is used for the evaluation of $\rho(\cdot)$ in (6). Then, $\hat{f}_i(x_i) = \hat{\mu}_i(x_i)$ and the following properties can be derived by considering the event-trigger (5) with (6).

Property 1: For any agent $i \in \mathcal{V}$ using the event-trigger (5) with (6), the following statements hold.
  a) If $c|\tilde{x}_i| \leq (\sqrt{N-1}+1)\underline{\eta}$, then $\hat{\eta}_i(x_i) \leq \underline{\eta}$.
  b) If $c|\tilde{x}_i| > (\sqrt{N-1}+1)\underline{\eta}$ and $\rho(x_i, \bar{x}_i) > 0$, then $\hat{\eta}_i(x_i) \leq \underline{\eta}$.
  c) If $c|\tilde{x}_i| > (\sqrt{N-1}+1)\underline{\eta}$ and $\rho(x_i, \bar{x}_i) \leq 0$, then
$$c^2|\tilde{x}_i|^2 \geq \eta_i^2(x_i) + (N-1)\underline{\eta}^2.$$

Proof: To prove the first statement, two cases with different sign of $\rho(\cdot)$ are considered. The case with $\rho(x_i, \bar{x}_i) > 0$ activates the data collection on agent i and then results in $\hat{\eta}_i(x_i) = \eta_i^+(x_i) \leq \underline{\eta}$ by comparing the posterior variance as in (8). For the case with $\rho(x_i, \bar{x}_i) \leq 0$, no model update occurs on agent i, and from the design of $\rho(\cdot)$ in (6) it holds

$$\hat{\eta}_i(x_i) = \eta_i(x_i) \leq c|\tilde{x}_i| - \sqrt{N-1}\underline{\eta} \leq \underline{\eta}. \qquad (9)$$

Combining the results from these two cases, the first statement is proven.

The second statement is directly obtained by considering that the model update is activated at agent i, since $\rho(x_i, \bar{x}_i) > 0$. Therefore, the boundness of the prediction error $\hat{\eta}_i(x_i) = \eta_i^+(x_i)$ by $\underline{\eta}$ is derived by using (8).

The proof of the third statement considers the non-positivity of $\rho(x_i, \bar{x}_i)$, where $\max\{c|\tilde{x}_i| - \sqrt{N-1}\underline{\eta}, \underline{\eta}\} = c|\tilde{x}_i| - \sqrt{N-1}\underline{\eta}$. Then, $\rho(x_i, \bar{x}_i) < 0$ is reformulated as

$$c^2|\tilde{x}_i|^2 \geq (\eta_i(x_i) + \sqrt{N-1}\underline{\eta})^2 \geq \eta_i^2(x_i) + (N-1)\underline{\eta}^2,$$

where the second inequality is derived by considering the positivity of $\eta_i(x_i)$ and $\underline{\eta}$ under $N \geq 1$. ∎

Remark 3: The maximum operator in (6) ensures the online learning only occurs when $\eta_i(x_i) > \underline{\eta}$, which saves the unnecessary data collection and induces high data efficiency.

Note that Property 1 only shows necessary statements for further derivation, i.e., the analysis of control performance, which is shown in Section III-C.

C. Consensus Performance Analysis

To analyze the control performance under distributed control law (3) with auxiliary dynamics (4) and event-triggered online learning (6), the error between the state x_i of agent i and their ideal average $\bar{x}^*$ defined in Definition 1 is studied, which is shown as follows.

Theorem 1: Consider a MAS containing N agents under a connected undirected topology $\mathcal{G}$ with individual dynamics (1) satisfying Assumption 1, which is controlled by (3) with (4). To infer the unknown function $f(\cdot)$ in (3), a GP model satisfying Assumption 3 is employed for each agent with a data set under Assumption 2. Moreover, each agent adopts the decentralized event-trigger (5) and (6) for online learning by adding new data pairs into the individual data set. Choose $\delta \in (0, 1/N)$, then the system achieves probabilistic ϵ-average consensus with $\epsilon = 2c^{-1}N\underline{\eta}$, i.e.,

$$\Pr\{\lim_{t \to \infty} \|\boldsymbol{x}(t) - \mathbf{1}_N \bar{x}^*\| \leq \epsilon\} \geq 1 - N\delta.$$

The proof of Theorem 1 is divided into two parts, namely the consensus of the auxiliary state (4) in Lemma 2 and the ultimate boundness of the consensus error in Lemma 3.

Lemma 2: Consider a multi-agent system with N agents under a connected undirected graph $\mathcal{G}$, whose auxiliary dynamics is designed as in (4), such that $\dot{\bar{\boldsymbol{x}}} = -\bar{c}\boldsymbol{\mathcal{L}}\bar{\boldsymbol{x}}$ with $\bar{\boldsymbol{x}} = [\bar{x}_1, \cdots, \bar{x}_N]^T$. Then, all the observed auxiliary states $\bar{x}_i$ for $\forall i \in \mathcal{V}$ converge to the average state $\bar{x}^*$ recalled from Definition 1 as $\bar{x}^* = N^{-1}\mathbf{1}_N^T\boldsymbol{x}(0)$, i.e., $\lim_{t \to \infty} \bar{\boldsymbol{x}}(t) = \mathbf{1}_N\bar{x}^*/N$.

Proof sketch: The proof of Lemma 2 follows [31] by showing the negative semi-definite of $\dot{V}_o$ for the positive definite Lyapunov function $V_o = \bar{\boldsymbol{x}}^T\bar{\boldsymbol{x}}$. Combining with the LaSalle's invariance principle, the asymptotic stability is shown with the equilibrium point of $\bar{x}_i$ as $\bar{x}^*$ by considering $\mathbf{1}_N^T\dot{\bar{\boldsymbol{x}}} = 0$ for the connected undirected graph. ∎

Remark 4: Considering the auxiliary state $\bar{x}_i$ is approaching the average state $\bar{x}^*$ due to (4), $\bar{x}_i$ can be regarded as the estimated average state for each agent.

Next, we prove the boundness of the consensus error related to $\tilde{\boldsymbol{x}} = [\tilde{x}_1, \cdots, \tilde{x}_N]^T$, which is shown as follows.

Lemma 3: Let all assumptions in Theorem 1 hold and choose $\delta \in (0, 1/N)$, then it holds $\lim_{t \to \infty} \|\tilde{\boldsymbol{x}}(t)\| \leq 2c^{-1}N\underline{\eta}$ with probability of at least $1 - N\delta$.

Proof: Apply the control law (3) with (4) into (1), the concatenated dynamics of the controlled system denotes

$$\dot{\tilde{\boldsymbol{x}}} = \dot{\boldsymbol{x}} - \dot{\bar{\boldsymbol{x}}} = -c(\boldsymbol{\mathcal{L}} + \boldsymbol{I}_N)\tilde{\boldsymbol{x}} + \boldsymbol{f}(\boldsymbol{x}) - \hat{\boldsymbol{\mu}}(\boldsymbol{x}), \qquad (10)$$

where the concatenated unknown function and its prediction denote $\boldsymbol{f}(\boldsymbol{x}) = [f(x_1), \cdots, f(x_N)]^T$ and $\hat{\boldsymbol{\mu}}(\boldsymbol{x}) = [\hat{\mu}_1(x_1), \cdots, \hat{\mu}_N(x_N)]^T$ respectively with $\hat{\mu}_i(\cdot)$ defined in Section III-B. Consider the controlled system with event-triggered learning as a switching system, and then the control performance is proven using the common Lyapunov theory. Choose the common Lyapunov candidate V_c as $V_c = \tilde{\boldsymbol{x}}^T \tilde{\boldsymbol{x}}$, whose time derivative using (10) is written as

$$\dot{V}_c = -2c\tilde{\boldsymbol{x}}^T(\boldsymbol{\mathcal{L}} + \boldsymbol{I}_N)\tilde{\boldsymbol{x}} + 2\tilde{\boldsymbol{x}}^T(\boldsymbol{f}(\boldsymbol{x}) - \hat{\boldsymbol{\mu}}(\boldsymbol{x})).$$

Consider $\underline{\lambda}(\boldsymbol{\mathcal{L}} + \boldsymbol{I}_N) = 1$ [31], where $\underline{\lambda}(\cdot)$ returns the minimal eigenvalue of the matrix. Then, $\dot{V}_c$ is bounded by

$$\dot{V}_c \leq -2c\|\tilde{\boldsymbol{x}}\|^2 + 2\|\tilde{\boldsymbol{x}}\|\|\boldsymbol{f}(\boldsymbol{x}) - \hat{\boldsymbol{\mu}}(\boldsymbol{x})\| \qquad (11)$$
$$\leq -2\|\tilde{\boldsymbol{x}}\|(c\|\tilde{\boldsymbol{x}}\| - \|\hat{\boldsymbol{\eta}}(\boldsymbol{x})\|)$$

with a probability of at least $1 - N\delta$ by using the union bound, where $\hat{\boldsymbol{\eta}}(\boldsymbol{x}) = [\hat{\eta}_1(x_1), \cdots, \hat{\eta}_N(x_n)]^T$. Note that the negativity of $\dot{V}_c$ is achieved when

$$c\|\tilde{\boldsymbol{x}}\| > \|\hat{\boldsymbol{\eta}}(\boldsymbol{x})\| \;\Leftrightarrow\; c^2\|\tilde{\boldsymbol{x}}\|^2 > \|\hat{\boldsymbol{\eta}}(\boldsymbol{x})\|^2.$$

Next, according to the value of $\tilde{x}_i$ and $\rho(x_i, \bar{x}_i)$, the agents are divided into 3 sets defined as

$$\mathbb{S}_1 = \{i \in \mathcal{V} : c|\tilde{x}_i| \leq (\sqrt{N-1}+1)\underline{\eta}\},$$
$$\mathbb{S}_2 = \{i \in \mathcal{V} : c|\tilde{x}_i| > (\sqrt{N-1}+1)\underline{\eta} \;\wedge\; \rho(x_i, \bar{x}_i) > 0\},$$
$$\mathbb{S}_3 = \{i \in \mathcal{V} : c|\tilde{x}_i| > (\sqrt{N-1}+1)\underline{\eta} \;\wedge\; \rho(x_i, \bar{x}_i) \leq 0\}.$$

satisfying $\bigcup_{i=1,2,3} \mathbb{S}_i = \mathcal{V}$ and $\mathbb{S}_i \cap \mathbb{S}_j = \emptyset$ for $i, j = 1, 2, 3$ and $i \neq j$. Considering $\sqrt{N} - \sqrt{N-1}$ is monotonically decreasing with respect to $N \in \mathbb{N}_+$, it has $\sqrt{N} - \sqrt{N-1} \leq \sqrt{1} - \sqrt{1-1}$ indicating $\sqrt{N} \leq \sqrt{N-1}+1$. This also means when $i \in \mathbb{S}_2 \cup \mathbb{S}_3$, it is easy to see $c|\tilde{x}_i| > \sqrt{N}\underline{\eta}$. Using the results in Property 1, the value of $\|\tilde{\boldsymbol{x}}\|^2$ is bounded as

$$c^2\|\tilde{\boldsymbol{x}}\|^2 = c^2 \sum_{j=1}^{3} \sum_{i \in \mathbb{S}_j} |\tilde{x}_i|^2 \geq \sum_{i \in \mathbb{S}_2} c^2|\tilde{x}_i|^2 + \sum_{i \in \mathbb{S}_3} c^2|\tilde{x}_i|^2$$
$$> \sum_{i \in \mathbb{S}_2} N\underline{\eta}^2 + \sum_{i \in \mathbb{S}_3} \left(\eta_i^2(x_i) + (N-1)\underline{\eta}^2\right), \quad (12)$$

where the first inequality is derived by considering $|\tilde{x}_i| \geq 0$, and the second inequality is due to the lower bound of $c|\tilde{x}_i|$ in Property 1 for sets $\mathbb{S}_2$ and $\mathbb{S}_3$. Moreover, considering $\eta_i(\cdot) = \hat{\eta}_i^2(\cdot)$ for agents $i \in \mathbb{S}_3$ without GP model update, the inequality in Eq. (12) is further written as

$$c^2\|\tilde{\boldsymbol{x}}\|^2 > \sum_{i \in \mathbb{S}_2} N\underline{\eta}^2 + \sum_{i \in \mathbb{S}_3} \left(\hat{\eta}_i^2(x_i) + (N-1)\underline{\eta}^2\right) \quad (13)$$
$$= |\mathbb{S}_2 \cup \mathbb{S}_3|N\underline{\eta}^2 + \sum_{i \in \mathbb{S}_3} \left(\hat{\eta}_i^2(x_i) - \underline{\eta}^2\right).$$

Similarly, the value of $\|\hat{\boldsymbol{\eta}}(\boldsymbol{x})\|^2$ is written as

$$\|\hat{\boldsymbol{\eta}}(\boldsymbol{x})\|^2 = \sum_{j=1}^{3} \sum_{i \in \mathbb{S}_j} \hat{\eta}_i^2(x_i) \leq \sum_{i \in \mathbb{S}_1 \cup \mathbb{S}_2} \underline{\eta}^2 + \sum_{i \in \mathbb{S}_3} \hat{\eta}_i^2(x_i)$$
$$= N\underline{\eta}^2 + \sum_{i \in \mathbb{S}_3} \left(\hat{\eta}_i^2(x_i) - \underline{\eta}^2\right). \quad (14)$$

Apply the inequalities (13) and (14) into (11), and then due to the positive c, (11) is bounded by

$$\dot{V}_c \leq -\frac{2\|\tilde{\boldsymbol{x}}\|}{c\|\tilde{\boldsymbol{x}}\| + \|\hat{\boldsymbol{\eta}}(\boldsymbol{x})\|}(c^2\|\tilde{\boldsymbol{x}}\|^2 - \|\hat{\boldsymbol{\eta}}(\boldsymbol{x})\|^2) \quad (15)$$
$$< -\frac{2\|\tilde{\boldsymbol{x}}\|}{c\|\tilde{\boldsymbol{x}}\| + \|\hat{\boldsymbol{\eta}}(\boldsymbol{x})\|}(|\mathbb{S}_2 \cup \mathbb{S}_3|N\underline{\eta}^2 - N\underline{\eta}^2).$$

Now, we consider the case with $\|\tilde{\boldsymbol{x}}\| > 2c^{-1}N\underline{\eta}$. This indicates at least one agent satisfying $|\tilde{x}_i| > 2c^{-1}\sqrt{N}\underline{\eta}$, which is proven by considering the contradiction with all agents having $c|\tilde{x}_i| \leq 2\sqrt{N}\underline{\eta}$ and

$$\|\tilde{\boldsymbol{x}}\|^2 \leq \sum_{i \in \mathcal{V}} 4c^{-2}N\underline{\eta}^2 = (2c^{-1}\sqrt{N}\underline{\eta})^2.$$

Note that $2\sqrt{N} \geq \sqrt{N-1}+1$ for $N \in \mathbb{N}$, therefore at least one agent belongs to $\mathbb{S}_2 \cup \mathbb{S}_3$. This also means when $\|\tilde{\boldsymbol{x}}\| > 2c^{-1}N\underline{\eta}$ the inequality $|\mathbb{S}_2 \cup \mathbb{S}_3| \geq 1$ holds, which is then applied in (15) such that $\dot{V}_c$ is bounded as

$$\dot{V}_c < -\frac{2\|\tilde{\boldsymbol{x}}\|N\underline{\eta}^2}{c\|\tilde{\boldsymbol{x}}\| + \|\hat{\boldsymbol{\eta}}(\boldsymbol{x})\|}(|\mathbb{S}_2 \cup \mathbb{S}_3| - 1) \leq 0.$$

This results in the decay of Lyapunov function V when $\|\tilde{\boldsymbol{x}}\| > 2c^{-1}N\underline{\eta}$, indicating $2c^{-1}N\underline{\eta}$ is the ultimate bound for $\|\tilde{\boldsymbol{x}}\|$, which concludes the proof. ∎

Remark 5: The proof for Lemma 3 also indicates the conventional control law (2) cannot achieve average consensus under non-vanishing prediction error $\boldsymbol{f}(\boldsymbol{x}) - \hat{\boldsymbol{\mu}}(\boldsymbol{x})$, since the minimal eigenvalue of $\boldsymbol{\mathcal{L}}$ is zero inducing infinite large ϵ by considering $\epsilon = (c\underline{\lambda}(\boldsymbol{\mathcal{L}}))^{-1}N\underline{\eta} \to \infty$ from (11) with $\boldsymbol{\mathcal{L}} + \boldsymbol{I}_N$ replaced by $\boldsymbol{\mathcal{L}}$.

Apply the intermediate result from Lemma 2 and Lemma 3, the proof of Theorem 1 is ready.

Proof of Theorem 1: Using the triangular inequality, the average consensus error $\|\boldsymbol{x} - \mathbf{1}_N \bar{x}^*\|$ is ultimately bounded by

$$\lim_{t \to \infty} \|\boldsymbol{x} - \mathbf{1}_N \bar{x}^*\| \leq \lim_{t \to \infty} \|\tilde{\boldsymbol{x}}\| + \lim_{t \to \infty} \|\bar{\boldsymbol{x}} - \mathbf{1}_N \bar{x}^*\|$$
$$= \lim_{t \to \infty} \|\tilde{\boldsymbol{x}}\| \leq 2c^{-1}N\underline{\eta} = \epsilon,$$

inherited the probability of at least $1 - N\delta$ from Lemma 3 by considering the probability of Lemma 2 as 1. ∎

Theorem 1 shows the ultimate boundedness of the consensus error with high probability, where the consensus error bound ϵ is increasing with the number of agents N. This relationship is intuitive, because in the worst case the prediction errors of the individual Gaussian process model on each body are superimposed, leading to a larger prediction error $\hat{\boldsymbol{\eta}}(\boldsymbol{x})$ with more agents. Moreover, the error ϵ is also related to the quality of the collected data reflected by $\underline{\eta}$, which is proportional to the variance of measurement noise σ_n considering Lemma 1 and (8). While using more accurate

sensors or processing methods reduces measurement noise, the variance σ_n will not vanish in practice. To achieve arbitrary small consensus error ϵ, a high gain c is necessary. The size of the ultimate bound ϵ only depends on the control gain c, while the auxiliary gain $\bar{c}$ determines the convergence speed [31]. This means larger $\bar{c}$ results in faster convergence, which is obtained by considering the convergence speed of the auxiliary state in Lemma 2 with [31] and the fact that Lemma 3 only considers $\tilde{x}_i = x_i - \bar{x}_i$. The decoupled effects of c and $\bar{c}$ provide us more flexibility to design the controller for desired performance in the steady state, i.e., for smaller ultimate average consensus errors, and the transition phase, i.e., for faster convergence, separately.

Remark 6: Note that only with $x_i(t) \in \mathbb{X}, \forall t \in \mathbb{R}_{0,+}$, the boundness of the prediction error in Lemma 1 and thus in Theorem 1 holds. Moreover, it is easily to prove the trajectories to $\bar{x}^*$ for each agent are in $\mathbb{X}$ due to the convex domain of $\mathbb{X}$ and $x_{(}0) \in \mathbb{X}$ for $\forall i \in \mathcal{V}$. Therefore, Theorem 1 holds with safety guarantee for the controlled system, if the neighbor domain of trajectories to $\bar{x}^*$ is a subset of $\mathbb{X}$. This can be easily realized by restraining the initial condition $x_i(0)$ to a smaller subset $\mathbb{X}_0 \subset \mathbb{X}$ such that the domain of $x(\infty)$, i.e., $\{x \in \mathbb{R} : |x - \bar{x}^*| \leq \epsilon\}$, is a subset of $\mathbb{X}$ with $\epsilon = 2c^{-1}N\bar{\eta}$ defined in Theorem 1 and $\bar{\eta} = \max_{i \in \mathcal{V}, x_i \in \mathbb{X}} \eta_i(x_i)$ from [14].

Remark 7: With the definition of ϵ in Theorem 1, the trigger condition in (6) can be relaxed as

$$\eta_i(x_i) > c^{-1} \max\{|\tilde{x}_i| - \epsilon/\sqrt{N}, 0\} + \underline{\eta},$$

which is composed of individual consensus error $\tilde{x}_i$, best prediction performance $\underline{\eta}$ and overall consensus error bound ϵ. This formulation indicates a general event-trigger design for online learning in multi-agent system, namely comparison of the current prediction error $\eta_i(x_i)$, the deviation of the control error to its bound $|\tilde{x}_i| - \epsilon/\sqrt{N}$, and the guaranteed best prediction performance $\underline{\eta}$, which is also suitable for cooperative learning with different hyperparameters for each Gaussian process model as in [14].

IV. NUMERICAL SIMULATIONS

A. Simulation Setting

To evaluate the designed control law (3), (4) with event-triggered online learning with (5), (6), we consider a multi-agent system with $N = 4$ agents under dynamics (1) with $h(x_i) = 0$, $g(x_i) = 1$ satisfying Assumption 1 and

$$f(x_i) = \sin(10x_i) + (2\exp(-x_i/10))^{-1} + 5$$

for each agent $i \in \mathcal{V} = \{1, \cdots, 4\}$ with the compact domain $\mathbb{X} = [-1.5, 1.5]$. The communication topology among the agents is defined by a connected undirected graph $\mathcal{G}$ with the edge set $\mathcal{E} = \{(1,2),(2,3),(3,4),(4,1)\}$. To infer the unknown function $f(\cdot)$, each agent $i \in \mathcal{V}$ employs a Gaussian process model using the kernel with Assumption 3 as

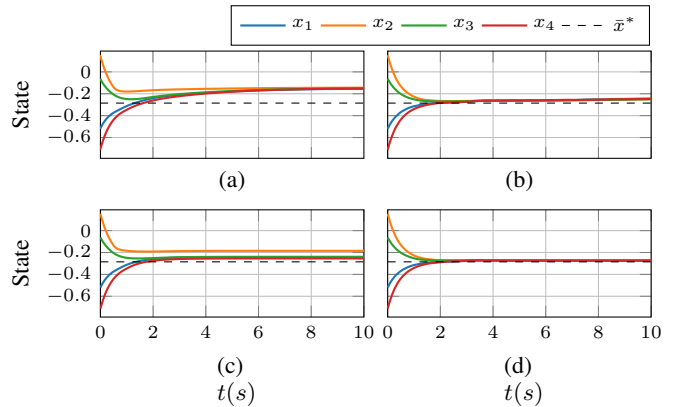$$\kappa(x, x') = \sigma_f^2 \exp\left(-0.5l^{-2}|x - x'|^2\right)$$

Fig. 1. The average state $\bar{x}^*$ and the actual state x_i for each agent over time using different controllers and learning strategies, which are corresponding to the 4 cases in Section IV-A.

with $\sigma_f = 1$, $l = 0.05$ and the individual data set $\mathbb{D}_i$ satisfying Assumption 2 with the variance of the measurement noise as $\sigma_n = 0.01$. The simulation time is set as 10.

To show the effectiveness of the proposed controller with event-triggered online learning, the following cases with different controllers and learning strategies are compared:

a) Use the conventional distributed control law (2) for average consensus with merely offline learning;
b) Use the conventional distributed control law (2) with event-triggered online learning using naive trigger function $\rho_{\text{naive}}(\cdot, \cdot)$ for all $i \in \mathcal{V}$ with

$$\rho_{\text{naive}}(x_i, \bar{x}_i) = \eta_i(x_i) - \underline{\eta}; \qquad (16)$$

c) Use the proposed controller (3) and (4) with merely offline learning;
d) Use the proposed controller (3) and (4) with event-triggered online learning using (5) and (6);

Note that the control gains c in (3) and (2) are set as $c = 1$, and $\bar{c} = 1$ for (4). Moreover, for online learning, the data sets $\mathbb{D}_i$ are initialized with 0 samples. For the offline learning cases, each agent employs an individual data set $\mathbb{D}_i$ including 150 training samples evenly distributed in the compact domain $\mathbb{X}$. Note that the number 150 for the size of the offline data set is chosen, such that the number of overall collected data using event-triggered online learning in cases (b) and (d) is below 150 with high probability, as shown in Fig. 4.

B. Performance among Difference Controllers

First, we show the performance for each cases with $x(0) = [-0.52, 0.15, -0.06, -0.71]^T$ resulting in $\bar{x}^* = -0.285$.

The state evolution for each agent over time is shown in Fig. 1. The performance by using the conventional controller (2) is shown in subfigures (a) and (b), where the states almost achieve consensus with some errors but do not converge to the average states $\bar{x}^*$. The bias between the actual converge state $\bar{x}_\infty = \lim_{t \to \infty} \mathbf{1}_N^T x(t)$ and the average states $\bar{x}^*$ can hardly be quantified. Therefore, it is not suitable for safety critical scenarios, since the actual converge state $\bar{x}_\infty$ may not in the pre-defined domain $\mathbb{X}$. Moreover, it is intuitive that with online learning the actual converge state is closer to the average states, because of the smaller prediction error
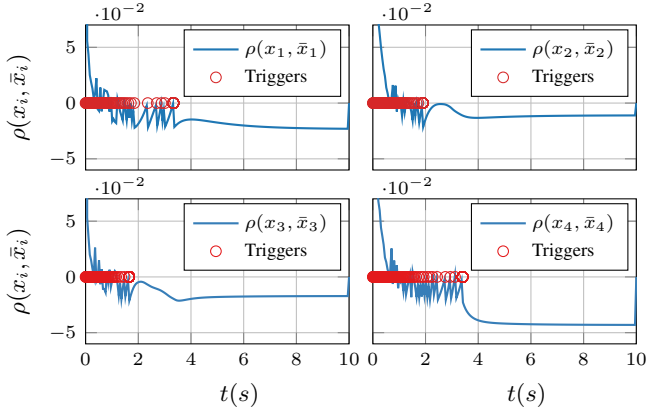
Fig. 2. The value of trigger function $\rho(x_i, \bar{x}_i)$ and the time of the trigger events for each agent $i \in \mathcal{V}$ for case (d).
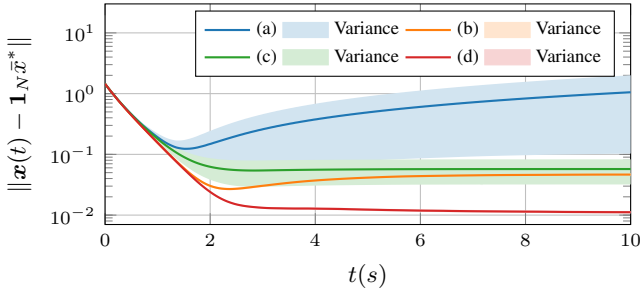

Fig. 3. Average consensus error over time with variance.

compared to the offline learning. With the proposed control using (3) and (4), the bias between $\bar{x}_\infty$ and $\bar{x}^*$ is guaranteed to be reduced and shown in subfigures (c) and (d) in Fig. 1, guarantee the effectiveness of Theorem 1 as in Remark 6. While the states x_i tend to close $\bar{x}^*$ in (c), the bias to $\bar{x}^*$, i.e., $|\bar{x}_\infty - \bar{x}^*|$, is obvious due to the poor prediction under offline learning. Additionally, with online learning as in (d), the controlled system achieves a smaller average consensus error compared to case (c).

Fig. 2 shows the event times for data collection for online learning with the values of trigger functions ρ in each agent. It is obvious that most data collections occur in the transition status as shown in Fig. 1 before $t = 4$, and in the steady states, i.e., $t > 4$ no trigger is activated indicating the Gaussian process models in each agent is sufficiently accurate to achieve the desired ϵ-average consensus. In particular, totally 82, 100, 71 and 112 trigger events occur on agent 1 to 4 respectively, resulting in smaller training data sets compared to the offline setting with $|\mathbb{D}_i| = 150$.

C. Monte Carlo Test

To show the generalization of the proposed controller with event-triggered learning, a Monte Carlo test is implemented, in which each case in Section IV-A is repeated 100 times with evenly randomly selected initial states $x_i(0) \in \mathbb{X}$ for $\forall i \in \mathcal{V}$ and random initial data sets for case (a) and (c).

The average consensus error $\|x - \mathbf{1}_N \bar{x}^*\|$ for each case is shown in Fig. 3, where the cases (a) and (c) with merely offline learning induce larger variance, showing strong dependency on the initial training sets. In comparison, two online learning approaches (b) and (d) demonstrate smaller
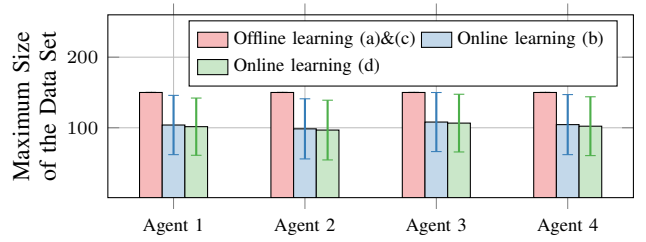

Fig. 4. Maximal size of data set, which for offline learning in cases (a) and (c) is 150 from the size of initial data set. With naive event-triggered online learning (16) in case (b), the maximal size of the data set for each agent 1 to 4 denotes 104 ± 42, 98 ± 43, 108 ± 42 and 104 ± 43, respectively. The maximal numbers of training samples in case (d) collected through decentralized event-triggered mechanism with (6) are 102 ± 41, 97 ± 42, 107 ± 41 and 102 ± 42 for agent 1 to 4, respectively.

average consensus error with smaller variance, because the training sets created through online data collection have desired prediction accuracy for the average consensus problem, i.e., $\hat{\eta}_i(\cdot) \leq \underline{\eta}$ as in (8). Moreover, there exists a divergence trend for two cases (a) and (b) with conventional distributed control law by considering the increasing average consensus error, indicating the observations in Fig. 1 (a) and (b) are common. Due to the larger bias between actual convergence state $\bar{x}_\infty$ and $\bar{x}^*$ by using (2), it is intuitive that cases (a) and (b) induce larger average consensus errors than case (c) and (d) with the proposed controller.

To show the comparison in Fig. 3 makes sense with similar number of training data, the size of the eventual data set in each agent is shown in Fig. 4. It is shown both event-triggered online learning strategies results in a data set with average near 100 samples and almost bounded by 150 samples. This indicates the predefined size of initial data set for cases (a) and (b) is sufficient large for data efficiency comparison with event-triggered online learning in cases (b) and (d). Combining with the results in Fig. 3, it is demonstrated that with similar number of data sets event-triggered online learning performs better in the control performance with smaller average consensus errors.

Until here, all results derived for the proposed distributed controller with decentralized event-triggered online learning in Section III are observed, indicating the effectiveness of the proposed methods.

V. CONCLUSION

In this paper, we propose an innovative approach to achieving average consensus in MASs considering unknown dynamics. This approach leverages a distributed consensus control law, which integrates an auxiliary dynamics. Additionally, to enhance control performance in real-time online learning scenarios and ensure high data efficiency, a decentralized event-trigger mechanism is designed for each agent. Remarkably, this mechanism operates solely on local state information, obviating the necessity for knowledge of the global communication topology. The theoretical proof of this work substantiates the stability of the system while guaranteeing an upper bound on the average consensus error making it particularly suitable for safety-critical applications. We conclude by demonstrating the effectiveness of our

proposed control and learning strategies through comparative numerical simulation.

APPENDIX

In the appendix, an example is given to show the states of the system (1) controlled by (2) are unbounded under bounded disturbance. We observe a simple MAS with only 2 agents under fully connected communication topology. Moreover, let the dynamics of each agent follows (1), which is then controlled by (2). Then, the concatenated controlled system is written as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = -c \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} f(x_1) - \mu_1(x_1) \\ f(x_2) - \mu_1(x_2) \end{bmatrix} \quad (17)$$

with $\mu_i(x_i)$ is sufficiently accurate such that only a small constant prediction error $\varepsilon \in \mathbb{R}_+$ is left, i.e., $f(x_i) - \mu_i(x_i) = \varepsilon$ for all $i = 1, 2$. The explicit solution of (17) denotes

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{x_1(0) + x_2(0) + 2\varepsilon t}{2} - \begin{bmatrix} 1 \\ -1 \end{bmatrix} \frac{x_1(0) - x_2(0)}{2} \exp(-2ct)$$

For non-zero prediction error ε, the unboundness of the state trajectory is obvious to see by considering

$$\lim_{t \to \infty} x_i(t) = \big(x_1(0) + x_2(0)\big)/2 + \varepsilon \lim_{t \to \infty} t \to \infty,$$

for $i = 1, 2$. Therefore, the conventional control law (2) can only stabilize the MAS with (1) for completely known dynamics, i.e., $f(x_i) = \mu_i(x_i)$, and not suitable in our setting. There also exists other robust or adaptive control laws for average consensus, which will be regarded as future work due to our main focus on the decentralized event-trigger design for online learning in multi-agent systems.

REFERENCES

[1] Z. Yan, Z. Yang, X. Pan, J. Zhou, and D. Wu, "Virtual Leader-based Path Tracking Control for Multi-UUV Considering Sampled-data Delays and Packet Losses," *Ocean Engineering*, vol. 216, p. 108065, 2020.

[2] L. Wan, X. Dai, T. Welfonder, E. Petrova, and P. Pauwels, "Semi-Automated Thermal Envelope Model Setup for Adaptive Model Predictive Control with Event-Triggered System Identification," in *Building Simulation Conference 2023*, 2023.

[3] Z. Yin, X. Dai, Z. Yang, D. Xiao, F. Menolascina, S. Hirche, and H. Zhao, "A Learning-based Online Controller Tuning Method for Electric Motors using Gaussian Processes," *Authorea Preprints*, 2024.

[4] 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*, pp. 1–5, 2019.

[5] A. Amirkhani and A. H. Barshooi, "Consensus in Multi-agent Systems: A Review," *Artificial Intelligence Review*, vol. 55, no. 5, pp. 3897–3935, 2022.

[6] Z.-G. Hou, L. Cheng, and M. Tan, "Decentralized Robust Adaptive Control for the Multiagent System Consensus Problem Using Neural Networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 3, pp. 636–647, 2009.

[7] M. Cai, Z. Xiang, and J. Guo, "Adaptive Finite-time Consensus Protocols for Multi-agent Systems by using Neural Networks," *IET Control Theory & Applications*, vol. 10, no. 4, pp. 371–380, 2016.

[8] Y. Lv, J. Zhou, G. Wen, X. Yu, and T. Huang, "Fully Distributed Adaptive NN-Based Consensus Protocol for Nonlinear MASs: An Attack-Free Approach," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 4, pp. 1561–1570, 2022.

[9] X. Jin, S. Lü, and J. Yu, "Adaptive NN-Based Consensus for a Class of Nonlinear Multiagent Systems With Actuator Faults and Faulty Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 8, pp. 3474–3486, 2022.

[10] 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 *Proceedings of The 5th Annual Learning for Dynamics and Control Conference*, vol. 211, pp. 445–457, PMLR, 15–16 Jun 2023.

[11] Z. Yang, X. Dai, A. Dubey, S. Hirche, and G. Hattab, "Whom to Trust? Elective Learning for Distributed Gaussian Process Regression," 2024.

[12] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning, Cambridge, Mass: MIT Press, 2006.

[13] T.-Y. Huang, X. Dai, S. Zhang, A. Capone, V. Todorovski, S. Sosnowski, and S. Hirche, "Learning-based prescribed-time safety for control of unknown systems with control barrier functions," *arXiv preprint arXiv:2403.08054*, 2024.

[14] 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 *2021 60th IEEE Conference on Decision and Control*, pp. 4406–4411, 2021.

[15] S. S. Kia, B. Van Scoy, J. Cortes, R. A. Freeman, K. M. Lynch, and S. Martinez, "Tutorial on Dynamic Average Consensus: The Problem, Its Applications, and the Algorithms," *IEEE Control Systems Magazine*, vol. 39, no. 3, pp. 40–72, 2019.

[16] A. Lederer, Z. Yang, J. Jiao, and S. Hirche, "Cooperative Control of Uncertain Multiagent Systems via Distributed Gaussian Processes," *IEEE Transactions on Automatic Control*, vol. 68, no. 5, pp. 3091–3098, 2023.

[17] 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*, pp. 7254–7259, 2022.

[18] X. Dai, Z. Yang, and S. Hirche, "Cooperative Online Learning for Multi-Agent System Control via Gaussian Processes with Event-Triggered Mechanism: Extended Version," 2024.

[19] R. Olfati-Saber and R. Murray, "Consensus Problems in Networks of Agents with Switching Topology and Time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[20] Z. Li, Z. Duan, and F. L. Lewis, "Distributed Robust Consensus Control of Multi-agent Systems with Heterogeneous Matching Uncertainties," *Automatica*, vol. 50, no. 3, pp. 883–889, 2014.

[21] S. R. Chowdhury and A. Gopalan, "On Kernelized Multi-armed Bandits," in *Proceedings of the 34th International Conference on Machine Learning*, vol. 70 of *Proceedings of Machine Learning Research*, pp. 844–853, PMLR, 06–11 Aug 2017.

[22] E. T. Maddalena, P. Scharnhorst, and C. N. Jones, "Deterministic Error Bounds for Kernel-based Learning Techniques under Bounded Noise," *Automatica*, vol. 134, p. 109896, 2021.

[23] A. Lederer, J. Umlauft, and S. Hirche, "Uniform Error Bounds for Gaussian Process Regression with Application to Safe Control," in *Advances in Neural Information Processing Systems*, vol. 32, Curran Associates, Inc., 2019.

[24] Z. Yin, X. Dai, Z. Yang, Y. Shen, G. Hattab, and H. Zhao, "Learning-based Control for PMSM Using Distributed Gaussian Processes with Optimal Aggregation Strategy," in *IECON 2023- 49th Annual Conference of the IEEE Industrial Electronics Society*, pp. 1–7, 2023.

[25] D. Yang, W. Ren, X. Liu, and W. Chen, "Decentralized Event-triggered Consensus for Linear Multi-agent Systems Under General Directed Graphs," *Automatica*, vol. 69, pp. 242–249, 2016.

[26] 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," 2024.

[27] Z. Yang, X. Dai, A. Dubey, S. Hirche, and G. Hattab, "Pri-GP: Prior-Aware Distributed Gaussian Process Regression," 2024.

[28] J. Umlauft and S. Hirche, "Feedback Linearization Based on Gaussian Processes With Event-Triggered Online Learning," *IEEE Transactions on Automatic Control*, vol. 65, no. 10, pp. 4154–4169, 2020.

[29] X. Dai, A. Lederer, Z. Yang, and S. Hirche, "Gaussian Process-Based Event-Triggered Online Learning with Computational Delays for Control of Unknown Systems: Extended Version," 2023.

[30] C. K. Williams and F. Vivarelli, "Upper and Lower Bounds on the Learning Curve for Gaussian Processes," *Machine Learning*, vol. 40, pp. 77–102, 2000.

[31] C. Godsil and G. F. Royle, *Algebraic Graph Theory*, vol. 207. Springer Science & Business Media, 2001.