Automatica 176 (2025) 112235

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Event-triggered learning-based control for output tracking with unknown cost functions☆

Jiliang Song [a], Dawei Shi [a,*], Shu-Xia Tang [b], Hao Yu [a], Yang Shi [c]

[a] School of Automation, Beijing Institute of Technology, Beijing, China
[b] Department of Mechanical Engineering, Texas Tech University, Lubbock, United States
[c] Department of Mechanical Engineering, University of Victoria, Victoria, Canada

ARTICLE INFO

Article history:
Received 21 February 2024
Received in revised form 11 November 2024
Accepted 2 February 2025
Available online xxxx

Keywords:
Network control system
Event-triggered control
Gaussian processes
Extremum seeking

ABSTRACT

In this paper, a two-layer event-triggered learning-based control framework is proposed to address extremum seeking problem in networked control systems with limited communication resources and unknown cost function. In this framework, the lower layer is an event-triggered controller to drive the output to track the given setpoints generated from the upper layer, where a learning-based optimizer is developed to approach the extremum of the unknown cost function. Specifically, in the lower layer, an event-triggered output controller, based on a high-gain extended state observer, is designed to tackle uncertainties and disturbances. In the upper layer, a nonparametric gradient model is established, and then the gradient descent method is applied to generate setpoints for the tracking control. The update of the learning and optimization process is determined by the tracking performance of the lower layer. The stability and Zeno-freeness of the proposed event-triggered controller is proved. Furthermore, the dependence of the convergence rate of the proposed learning-based extremum seeking algorithm on the designed parameters is also explicitly characterized. Finally, the effectiveness of the proposed framework is validated by numerical examples.

© 2025 Published by Elsevier Ltd.

1. Introduction

In recent decades, the development of network technology has significantly advanced the field of networked control system (NCS), which is now increasingly favored due to the low maintenance cost and scalability (Heemels et al., 2010; Léchappé et al., 2019; Nesic & Teel, 2004; Sun et al., 2021). In some NCSs, the optimal reference signal is unknown or hard to determine, e.g., spacecraft antenna adjusting to an unknown source (Walsh et al., 2017). Such scenarios necessitate a controller capable of seeking the direction of the unknown source and ensuring the rapid tracking of the source while operating within the constraints of limited on-board communication resources.

Event-triggered control (ETC) has emerged as a promising solution to reduce the communication demands of NCSs, thereby improving resource utilization efficiency. This is primarily due to its non-periodic sampling characteristics, dictated by predefined event-triggering conditions (Heemels et al., 2012). Within an ETC framework, the event-triggering condition is pivotal in regulating event occurrences and, consequently, the overall control performance (Girard, 2015; Heemels & Donkers, 2013; Huang et al., 2020; Li & Liu, 2020; Sun et al., 2022; Xie et al., 2023). The ETM significantly influences ETC system performance by controlling sensor update time instants, affecting control execution, and minimizing unnecessary events, thus requiring a case-specific design approach. Despite the potential of a well-crafted ETM to maintain satisfactory control performance, an inherent trade-off between this efficacy and transmission frequency persists in ETC systems. This challenge is exacerbated under the existence of uncertainties and disturbances, especially with limited plant information. Strategies to counteract internal uncertainties and external disturbances generally fall into two categories: robust controller design (Abdelrahim et al., 2017; Li & Shi, 2014; Liu & Jiang, 2015) and disturbance observation (Li et al., 2020; Sun et al., 2022). Notably, the extended state observer (ESO) (Gao, 2003; Guo & Zhao, 2011b; Ran et al., 2021), stands out as a model-free method for estimating the disturbances and uncertainties, without requiring detailed plant information.

On the other hand, extremum-seeking control (ESC) offers a general framework for model-free optimization of unknown performance functions (Guay et al., 2004; Krstić & Wang, 2000;

☆ This work is supported by the National Natural Science Foundation of China under Grants 62261160575 and 62333001. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Tiago Roux Oliveira under the direction of Editor Miroslav Krstic.

* Corresponding author.
E-mail addresses: songjiliang@bit.edu.cn (J. Song), daweishi@bit.edu.cn (D. Shi), shuxia.tang@ttu.edu (S.-X. Tang), yuhaocsc@bit.edu.cn (H. Yu), yshi@uvic.ca (Y. Shi).

https://doi.org/10.1016/j.automatica.2025.112235
0005-1098/© 2025 Published by Elsevier Ltd.

Suttner & Krstić, 2023). Traditional ESC methods typically overlap a low-frequency dither signal onto the control input of the system to estimate the gradient online by extracting the variations of the performance function. However, the slow dither signal for the control is undesirable for some practical systems since it brings extra difficulties to the implementation of the controller. Additionally, the dither approach suffers from the timescale separation (Krishnamoorthy & Skogestad, 2022). To overcome this problem, Zhang and Ordóñez (2007) introduced a numerical optimization-based ESC method based on the line search and trust region methods. While the approach proposed by Zhang and Ordóñez (2009) demanded precise gradient measurement which may be unrealistic in practical applications, subsequent work by Ye and Hu (2016) adopted the methodology to scenarios with estimated gradients, providing bounds on gradient estimation errors to ensure convergence.

However, conventional gradient-estimation methods in ESC typically neglect measurement noise, which may lead to unacceptable deviations of the estimated gradient (Fu & Özgüner, 2011). To address the uncertainty inherent in noisy black-box models, the Bayesian framework, particularly Gaussian Process (GP), offers a robust alternative with favorable theoretical properties (Rasmussen & Williams, 2006). An important application of GP is Bayesian optimization (Srinivas et al., 2012), however, high-dimensional models present the curse of dimensionality, which has led to the development of local Bayesian optimization (BO) (Müller et al., 2021; Wu et al., 2023). In local BO, the Jacobian GP is utilized for gradient estimation while searching for the optimum by locally inquiring new samples and incrementally updating the Jacobian GP model. The philosophy of local BO inspires us to leverage online learning to address the extremum seeking problem with measurement noises. Note that optimizing output performance in ETC has not been adequately addressed, especially in scenarios with unknown plant models and performance functions, along with external disturbances. Thus, the convergence of a Jacobian GP-based extremum-seeking strategy and the stability of the closed-loop ETC system remain open issues, which are to be addressed in this work.

In this work, we develop a two-layer event-triggered learning-enabled tracking control framework for a class of NCSs, which integrates the event-triggered output tracking control with extremum seeking optimization for unknown performance function. Different from the dither-based approach for extremum seeking (Rodrigues et al., 2023; Zhu et al., 2023), which inserted a dither signal into the control input to obtain the continuous gradient estimation, the developed method in this work achieves extremum seeking through a numerical optimization-based approach (Zhang & Ordóñez, 2007, 2009) which only evaluates the gradient at specific positions. Furthermore, this work considers the presence of Gaussian noise in the measurement of the performance function and emphasizes the role of learning, optimization, and ETC strategies in supporting the stability of the proposed extremum-seeking framework.

To this end, the key challenges addressed in this paper are twofold: (i) The training data and hyperparameters are key determinants of the learning performance for GP, and it is challenging to construct a quantitative relationship between these factors and the gradient estimation error; (ii) convergence analysis of the upper and lower layers is closely related and non-trivial, especially when incorporating the event-triggered mechanisms: the convergence of the learning algorithm depends on the results of the control, and conversely the stability of the controller depends on the outputs of the upper learning algorithm. The main contributions of this work can be summarized as follows:

1. A two-layer event-triggered extremum seeking control framework with the GP learning is proposed. Specifically, a Jacobian GP-based extremum seeking optimizer provides a set reference trajectories for the output tracking controllers, and the data generated during the control process is also used to update the Jacobian GP model. Given this, the optimizer and controller are set to be asynchronous, and the optimizer will only update and provide the next reference point when the tracking error of the controller converges to the preset invariant set. In addition, due to the time-varying and discontinuous set points provided by the optimizer, we use a tracking differentiator to generate a continuously differentiable reference trajectory for the controller.
2. The convergence analysis associating with the gradient estimation error and setpoint tracking error for the learning-based extremum seeking control algorithm is provided. In particular, a data-and-hyperparameter dependent asymptotic bound for the gradient estimation error using Jacobian GP is obtained. We show that the upper bound of the gradient estimation error can be guaranteed to satisfy the convergence conditions of the extremum seeking optimization when the samples are collected within a compact set defined by the upper bound and the minimum inter-sample distance conditions are satisfied.
3. We synthesize the convergence analysis of the learning-based extremum seeking control in combination with the event-triggered output tracking control in the lower layer. Two event-triggering mechanisms are introduced in the sensor–observer and control–actuator channels, respectively, and the semi-global asymptotic stability and Zeno-freeness is proved. In particular, we show that by appropriately designing the event-triggered threshold for sampling at the output of the system, the samples can be automatically made to satisfy the minimum inter-data distance conditions for the learning algorithm in the upper layer to converge. Finally, the effectiveness of the proposed method is validated through numerical examples.

Notation. We denote $|\cdot|$ as the absolute value of a scalar and $\|\cdot\|$ as the Euclidean norm of a vector. For a matrix A, $\|A\|_F$ denotes the Frobenius norm. In particular, for special norms, we use the subscript to indicate the specific norms defined in special spaces, e.g., $\|J\|_{\mathcal{H}}$ in reproducing kernel Hilbert space. For a compact set $\mathcal{A}$, we denote $\partial\mathcal{A}$ as the boundary of $\mathcal{A}$ and $\text{int}(\mathcal{A})$ as the interior of $\mathcal{A}$. For a matrix B, we use $\bar{\lambda}(B)$ and $\underline{\lambda}(B)$ represent the maximum and minimum eigenvalues of B, respectively. In addition, $\text{tr}(B)$ represents the trace of B. We denote $[B]_i$ as the ith row of the matrix B. For a function $f(\cdot)$, we denote $\nabla f(\cdot)$ and $\nabla^2 f(\cdot)$ as the gradient and Hessian of $f(\cdot)$, respectively.

2. Problem formulation

In this work, we consider a class of strictly feedback nonlinear networked systems with external disturbances and unmodeled dynamics. For all $i \in \{1, \ldots, m\}$, the formulation of each ith nonlinear subsystem is described as follows:

$$\begin{aligned} \dot{x}_i(t) &= Ax_i(t) + B[f_i(t, x_i(t), w_i(t)) + u_i(t)], \\ y_i(t) &= x_{i1}(t), \end{aligned} \tag{1}$$

where $x_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^{\mathrm{T}} \in \mathbb{R}^n$ and $y_i \in \mathbb{R}$ are the states and output of the ith subsystem, $A = [0\ I_{n-1}; 0\ 0]_{n\times n}$, $B = [0, 0, \ldots, 0, 1]^{\mathrm{T}}_{n\times 1}$, the nonlinear function $f_i(\cdot) : [0, \infty) \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is unknown but continuously differentiable, $w_i(t) \in \mathbb{R}$ represents the unknown disturbance, and $u_i(t) \in \mathbb{R}$ is the control input.

For system (1), we have the following assumptions:

Assumption 1. For $i = 1, \ldots, m$, there exist continuous functions $\kappa_i(\cdot) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^+$ such that $\max_{(t,x_i,w_i)\in\mathbb{R}^+\times\mathbb{R}^n\times\mathbb{R}} \{|f(t, x_i, w_i)|, \|\nabla f(t, x_i, w_i)\|\} \leq \kappa_i(x_i, w_i)$.

Assumption 2. The disturbances w_i are unknown but bounded and twice continuous differentiable, satisfying $|w_i| + |\dot{w}_i| \leq c_w, i = 1, \ldots, m$.

Assumptions 1 and 2 are some classical assumptions on the unknown dynamics and disturbance which ensure that the system is analytically tractable and practical feasible (Guo & Zhao, 2011b; Ran et al., 2021). It is noteworthy that we consider the networked control system, namely, the nonlinear plants and controller are located separately while the output measurements and control signals are transmitted through communication networks.

In addition, we consider a global unknown, yet measurable (with noises), performance function, which is associated with the outputs of all subsystems:

$$\theta = J(y) + \upsilon, \tag{2}$$

where $\theta \in \mathbb{R}$, $y = [y_1, \ldots, y_m]$, and $\upsilon \in \mathbb{R}$ is Gaussian noise with zero mean and standard deviation $\sigma_g \in \mathbb{R}$. For this unknown function, we have the following assumptions.

Assumption 3. There exists an isolated global minimizer y^* satisfying that for all $y \neq y^*$, $J(y) > J(y^*) > -\infty$.

Definition 1 (*Smoothness*)**.** A function J is β-smooth if and only if for all $y_1, y_2 \in \mathcal{Y}$, where $\mathcal{Y}$ is a compact set with proper dimension, we have $\|\nabla J(y_1) - \nabla J(y_2)\| \leq \beta\|y_1 - y_2\|$.

Assumption 4. The performance function $J(y)$ is a sample from a Gaussian process with known mean, covariance kernel, and hyperparameters, and for $0 < \delta < 1$, a constant β exists such that with probability at least $1 - \delta$, $J(y)$ is β-smooth.

The control objective of this work is to regulate the output of each subsystem in (1) to the positions at which the minimum value of unknown performance function $J(y)$ is reached. We consider each subsystem to be independent of each other. To this end, the control strategy should be capable of optimizing the unknown performance with noisy measurements and achieving fine control of the unmodeled nonlinear systems.

3. Main results

In this section, a two-layer event-triggered learning-based extremum seeking control framework is proposed (see Fig. 1). The stability analysis of the proposed framework is then provided: we show the convergence of the learning-based extremum seeking algorithm as well as the stability of the event-triggered controller.

3.1. Jacobian GP-based extremum seeking

Suppose that a dataset $\mathcal{D}_0 = \{Y_0, \Theta_0\}$ can be gathered near the initial position at the outset, intended to initialize a Jacobian GP. The initial dataset $\mathcal{D}_0$ includes both the system outputs $Y_0 = \{y^{\{1\}}, y^{\{2\}}, \ldots, y^{\{d\}}\} \in \mathbb{R}^{m\times d}$ and noisy measurements $\Theta_0 = \{\theta^{\{1\}}, \theta^{\{2\}}, \ldots, \theta^{\{d\}}\}^{\mathrm{T}} \in \mathbb{R}^d$ of the unknown cost function. The superscript $\{i\}$ represents the ith elements of the dataset and $d \in \mathbb{N}^+$. Assuming $J(y)$ is GP-distributed and the kernel is differentiable, the gradient process $\nabla J(y)$ is also jointly GP-distributed (Rasmussen & Williams, 2006). Thus, for a test input $y^* \in \mathbb{R}^m$, the noisy observations Θ_0 and the gradient $\nabla J(y^*)$ are jointly Gaussian distributed as:

$$\begin{pmatrix} \Theta_0 \\ \nabla J(y^*) \end{pmatrix} \sim \mathcal{GP}\left(\begin{pmatrix} \mu(Y_0) \\ \nabla\mu(y^*) \end{pmatrix}, \begin{pmatrix} K(Y_0,Y_0)+\sigma_g^2 I & K(Y_0,y^*)\nabla^{\mathrm{T}} \\ \nabla K(y^*,Y_0) & \nabla K(y^*,y^*)\nabla^T \end{pmatrix}\right) \tag{3}$$

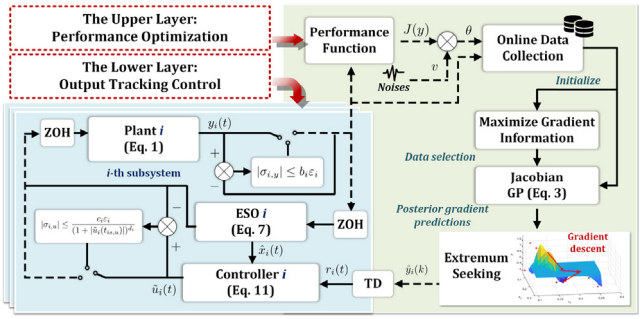


Fig. 1. Schematic of the event-triggered learning-based extremum seeking control framework. The solid line denotes the continuous signal and the dashed line denotes the discrete signal.

where $\mu(Y_0)$ and $\nabla\mu(y^*)$ represent the mean function and its gradient at the test input, respectively, and $K(\cdot,\cdot)$ is the covariance matrix defined by the kernel function $k(\cdot,\cdot) : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$. The posterior mean $\nabla\mu_0$ and covariance $\nabla\sigma_0^2$ of the gradient can be derived from the conditional joint Gaussian distribution $p(\nabla J(y^*)|y^*, Y_0, \Theta_0) \sim \mathcal{GP}(\nabla\mu_0, \nabla\sigma_0^2)$ with $\nabla\mu_0 = \nabla\mu(y^*) + \nabla K(y^*, Y_0)(K(Y_0, Y_0) + \sigma_g^2 I)^{-1}(\Theta_0 - \mu(Y_0))$ and $\nabla\sigma_0^2 = \nabla^2 K(y^*, y^*) - \nabla K(y^*, Y_0)(K(Y_0, Y_0) + \sigma_g^2 I)^{-1}\nabla K(Y_0, y^*)$. Generally, the posterior mean is taken as the estimation of the gradient, which can be denoted as $\nabla\mu_k$ for the kth step. Thus, the gradient descent strategy can be adopted with $\hat{y}_{k+1} = y_k - \alpha\nabla\mu_k$, where $\alpha > 0$ is the search step. Considering the tracking error of the controller, the finally reached y_{k+1} is biased with respect to the setpoint $\hat{y}_{k+1}$, i.e., $y_{k+1} = \hat{y}_{k+1} + \Lambda_k$, where $\Lambda_k \in \mathbb{R}^m$ is the control error.

In this work, we consider the square exponential (SE) kernel:

$$k_{SE}(y_1, y_2) = \theta_0^2 \exp(-\tfrac{1}{2\theta_1^2}\|y_1 - y_2\|^2), \tag{4}$$

where θ_0, $\theta_1 > 0$ are hyperparameters. Since the SE kernel is stationary, bounded, continuously differentiable, and strictly positive, we have $k_{SE}(y, y^*) = \phi_{SE}(y - y^*)$, $\nabla k_{SE}(y, y^*) = \nabla\phi_{SE}(y - y^*)$, and $\nabla k_{SE}(y, y^*)\nabla^{\mathrm{T}} = -\nabla^2\phi_{SE}(y - y^*)$ where $\phi_{SE}(\cdot)$ is a positive-definite function. Considering the SE kernel in (4), we have the following assumption on the function $J(y)$:

Assumption 5. The unknown function $J(y)$ has a bounded reproducing kernel Hilbert space (RKHS) norm w.r.t. the SE kernel on any compact set.

Assumption 5 characterizes the smoothness of the function $J(y)$ in the RKHS when considering the SE kernel, which can be leverage to characterize the approximation error of the GP, and is a common and standard assumption in the literature (Beckers et al., 2019; Srinivas et al., 2012; Wu et al., 2023).

Additionally, to reduce the computation time and enhance the learning efficiency, we adopt an active data selection strategy based on the posterior covariance of the test data. Denote the current position as $y_k = y(t_k)$ and the training dataset at the $(k-1)$-th step as $\mathcal{D}_{k-1} = \{Y_{k-1}, \Theta_{k-1}\}$. Let $\mathcal{Y}_k$ be a compact set near the position at the kth step and $\tilde{Y}_k \in \mathbb{R}^{m\times d_k}$ be the candidate test inputs collect from $\mathcal{Y}_k$. Then, we have the posterior covariance matrix conditioned on the training samples Y_{k-1} and the candidate test inputs $\tilde{Y}_k$ as $K_{Y_{k-1}\cup\tilde{Y}_k}(y_k, y_k) = K_{Y_{k-1}}(y_k, y_k) - K_{Y_{k-1}}(y_k, \tilde{Y}_k)(K_{Y_{k-1}}(\tilde{Y}_k, \tilde{Y}_k) + \sigma_g^2 I)^{-1}K_{Y_{k-1}}(\tilde{Y}_k, y_k)$. Hence, we have the following data selection strategy to maximize the gradient information (Müller et al., 2021).

$$\arg\min_{\tilde{Y}_k\subset\mathcal{Y}_k} \operatorname{tr}(\nabla K_{Y_{k-1}\cup\tilde{Y}_k}(y_k, y_k)\nabla^T). \tag{5}$$

By solving (5), a batch of samples $\tilde{Y}_k$ is selected from the compact set $\mathcal{Y}_k$ that minimize the one-step look-ahead uncertainty, as measured by the trace of the posterior covariance matrix.

Algorithm 1 briefly summarizes the main steps of the proposed Jacobian GP-based extremum seeking algorithm.

Algorithm 1 The Jacobian GP-based extremum seeking algorithm.

Initialize: $\mathcal{D}_0 = \{Y_0, \Theta_0\}$, $k = 0$, $y_0 = y(t_0)$.
while $k \geq 1$ and $\|\nabla\mu_k\|^2 \geq M_{grad}$ **do**
  Calculate μ_k;
  Apply gradient descent $\mathring{y}_{k+1} = y_k - \alpha\nabla\mu_k$;
  while $\|\Lambda_k\| \geq \|\nabla J(y_k)\|/10\beta$ **do**
    Control the output $y(t)$ to track $\mathring{y}_{k+1}$;
  end while
  Collect data $\mathcal{D}'_k$ from the neighborhood $\mathcal{Y}_k$ of the current position y_k and the corresponding function evaluations following the trajectories;
  Select d_k samples from $\mathcal{D}'_k$ to construct the new dataset $\tilde{\mathcal{D}}_k$ according to (5);
  $\mathcal{D}_{k+1} = \mathcal{D}_k \cup \tilde{\mathcal{D}}_k$ and update the GP model;
  $k = k + 1$.
end while

As shown in Algorithm 1, the training dataset $\mathcal{D}_k$ is composed of the last dataset $\mathcal{D}_{k-1}$ and the new acquired samples $\tilde{\mathcal{D}}_k = (\tilde{Y}_k, \tilde{\Theta}_k)$. Denote the number of samples in $\mathcal{D}_k$, $\mathcal{D}_{k-1}$, and $\tilde{\mathcal{D}}_k$ as N_k, N_{k-1}, and d_k, respectively. In the following lemma, we provide a uniform upper bound for the gradient estimation error at the kth step:

Lemma 1. *Consider an unknown function $J(y)$ satisfying Assumptions 4–5, a Jacobian GP in (3), and the SE kernel in (4). Then, for any $\delta \in (0,1)$ and all the training samples Y_k, the estimated gradient of the kth step satisfies* $\mathrm{P}\left\{\|\nabla J(y_k) - \nabla\mu_k\|^2 \leq C_k\Gamma_k\right\} \geq 1-\delta$, *where* $C_k = 2\log\left(\frac{\pi^2k^2}{6\delta}\right)$, $N_k = N_{k-1} + d_k$, $\Gamma_k = \frac{m\theta_0^2}{\theta_1^2} - \frac{\|y_k - y'\|^2 k_{SE}(y_k,y')^2}{\theta_1^4(k_{SE}(y'',y''')+\sigma_g^2/N_k)}$, $y' = \arg\min_{y'\in Y_k}\|y_k - y'\|^2 k_{SE}(y_k,y')^2$, $y'', y''' = \arg\max_{y'',y'''\in Y_k} k_{SE}(y'',y''')$.

Proof. We firstly present the following result (Wu et al., 2023, Lemma 11): For any $\delta \in (0,1)$, let $C_k = 2\log\left(\frac{\pi^2k^2}{6\delta}\right)$, then the following inequalities

$$\|\nabla J(y_k) - \nabla\mu_k\|^2 \leq C_k \mathrm{tr}(\nabla k(y_k,y_k)\nabla^{\mathrm{T}}) \tag{6}$$

hold for any $k > 1$ with probability at least $1-\delta$.

By Assumption 4, since $J(y_k)$ is a sample from GP, $\nabla J(y_k)$ is a sample from GP, which satisfies $\nabla J(y_k) \sim \mathcal{GP}(\nabla\mu_k, \nabla k(y_k,y_k)\nabla^{\mathrm{T}})$. Then, by the definition of the posterior covariance, we have $\mathrm{tr}(\nabla k(y_k,y_k)\nabla^{\mathrm{T}}) = \mathrm{tr}[\nabla^2 k(y_k,y_k) - \nabla k(y_k,Y_k)A_N^{-1}\nabla k(Y_k,y_k)]$ where $A_N = K(Y_k,Y_k) + \sigma_g^2 I$.

For the SE kernel in (4), let $N_k = d_0 + kd_k$, we can directly calculate that $\nabla^2 k_{SE}(y_k,y_k) = \mathrm{diag}\left(\frac{\theta_0^2}{\theta_1^2}\cdots\frac{\theta_0^2}{\theta_1^2}\right)_{m\times m}$. Meanwhile, we have $k_{SE}(y_k,Y_k) = \nabla^{\mathrm{T}} k_{SE}(Y_k,y_k) \in \mathbb{R}^{m\times N}$ with the (i,j)-th element as $\partial k_{SE}(y_k,Y_k(j))/\partial y_k(i)$ for all $i = 1\ldots m, j = 1\ldots N$. Since the covariance matrix $K(Y_k,Y_k)$ is positive semi-definite and the noise standard deviation $\sigma_g > 0$, the matrix A_N is symmetric and positive definite, then the maximum eigenvalue of A_N satisfies $\bar{\lambda}(A_N) = \bar{\lambda}(K(Y_k,Y_k)) + \sigma_g^2$. Then, we have $\underline{\lambda}(A_N^{-1}) = 1/(\bar{\lambda}(K(Y_k,Y_k)) + \sigma_g^2)$. Hence, it can directly obtained that $\mathrm{tr}\{\nabla k_{SE}(y_k,Y_k)A_N^{-1}\nabla k_{SE}(Y_k,y_k)\} \geq \frac{\|\nabla k_{SE}(y_k,Y_k)\|_F^2}{\bar{\lambda}(K(Y_k,Y_k))+\sigma_g^2}$. Subsequently, by defining $y' = \arg\min_{y'\in Y_k}\|y_k - y'\|^2 k_{SE}(y_k,y')^2$, we have $\|\nabla k_{SE}(y_k,Y_k)\|_F^2 \geq \frac{N}{\theta_1^4}\|y_k - y'\|^2 k_{SE}(y_k,y')^2$. In addition, according to Gerschgorin theorem, the maximum eigenvalue of $K(Y_k,Y_k)$ could be upper bounded by $\bar{\lambda}(K(Y_k,Y_k)) \leq a_{ii} + \sum_{j=1,j\neq i}^{N-1}|a_{ij}| \leq Nk(y'',y''')$, where a_{ij}'s are the entries of the covariance matrix $K(Y_k,Y_k)$ in the ith row, jth column, and $y'', y''' = \arg\max_{y'',y'''\in Y_k} k(y'',y''')$. Hence, we can calculate that $\mathrm{tr}(\nabla k_{SE}(y_k,y_k)\nabla^{\mathrm{T}}) \leq \frac{m\theta_0^2}{\theta_1^2} - \frac{\|y_k-y'\|^2 k_{SE}(y_k,y')^2}{\theta_1^4(k_{SE}(y'',y''')+\sigma_g^2/N_k)}$. Combining the above discussions with (6), the proof of Lemma 1 is completed. □

Lemma 1 establishes the relationships among the gradient estimation error, the SE kernel, and the training data. Based on Lemma 1, we have the following convergence result:

Theorem 1. *Consider Algorithm 1 and suppose Assumptions 1–5 hold. Define $\mathcal{Y}_k = \{y \in \mathbb{R}^m | \rho_1 \leq \|y_k - y\| \leq \rho_2\}$, where $0 \leq \rho_1 \leq \theta_1\sqrt{W(\frac{\xi_k}{\theta_0^4\theta_1^2})}$ and $\theta_1 \geq \rho_2 \geq \rho_1$. Let the training samples collected from $\mathcal{Y}_k$ satisfy $d_k \geq m+1$ and $\min_{y^{\{i\}},y^{\{j\}}\in\mathcal{D}_k, i\neq j}\|y^{\{i\}} - y^{\{j\}}\|^2 \leq -2\theta_1^2\ln(\epsilon)$ with $1/m > \epsilon > 0$, the hyperparameters θ_0, θ_1 satisfy $\theta_0^2 \leq \frac{m(1+(N_k-1)\epsilon)+\sqrt{m^2(1+(N_k-1)\epsilon)^2+4N_k m\sigma_g^2}}{2N_k}$ and $\theta_1^2 \geq \frac{25mC_k\theta_0^2(\theta_0^2+(N_k-1)\epsilon+\sigma_g^2)-25\theta_0^6N_kC_k}{\|\nabla J(y_k)\|^2(\theta_0^2+(N_k-1)\epsilon+\sigma_g^2)}$, and the search step satisfies $\alpha \in (\frac{1}{4\beta}, \frac{11}{36\beta})$. Then, for all $\delta \in (0,1)$, if the control error $\|\Lambda_k\| \leq \|\nabla J(y_k)\|/10\beta$, $J(y_{k+1}) - J(y_k) \leq 0$ with probability at least $1-\delta$.*

Proof. According to the descent lemma (Zhang & Ordóñez, 2007) and Assumption 4, we have $J(y_{k+1}) - J(y_k) \leq \nabla J(y_k)^{\mathrm{T}}(y_{k+1} - y_k) + \frac{1}{2}\beta\|y_{k+1} - y_k\|^2$ holds with probability $1-\delta$. Since the optimization is proceeded with bounded output tracking error at each step, i.e., $y_{k+1} = \mathring{y}_{k+1} + \Lambda_k = y_k - \alpha\nabla\mu_k + \Lambda_k$, we can obtain that $J(y_{k+1}) - J(y_k) \leq -\alpha\nabla J(y_k)^{\mathrm{T}}\nabla\mu_k + \nabla J(y_k)^{\mathrm{T}}\Lambda_k + \frac{\beta}{2}\|-\alpha\nabla\mu(y_k) + \Lambda_k\|^2 \leq -\alpha\|\nabla J(y_k)\|\|\nabla\mu_k\|\cos\varphi_k + \|\nabla J(y_k)\|\|\Lambda_k\| + \beta\alpha^2\|\nabla\mu_k\|^2 + \beta\|\Lambda_k\|^2$ hold with probability $1-\delta$, where φ_k is the angle between the vectors $\nabla J(y_k)$ and $\nabla\mu_k$ at the kth step.

By Lemma 1, we can obtain $\nabla J(y_k)^{\mathrm{T}}\nabla\mu_k = \nabla J(y_k)^{\mathrm{T}}(\nabla J(y_k) - (\nabla J(y_k) - \nabla\mu_k)) \geq \|\nabla J(y_k)\|(\|\nabla J(y_k)\| - \sqrt{C_k\Gamma_k})$ hold with probability at least $1-\delta$. For brevity of discussion, the explicit mention of probability is omitted in the following process of the proof. Then, $\cos\varphi_k = \frac{\nabla J(y_k)^{\mathrm{T}}\nabla\mu_k}{\|\nabla J(y_k)\|\|\nabla\mu_k\|} \geq \frac{\|\nabla J(y_k)\| - \sqrt{C_k\Gamma_k}}{\|\nabla\mu_k\|}$ and $-\alpha\|\nabla J(y_k)\|\|\nabla\mu_k\|\cos\varphi_k \leq -\alpha\|\nabla J(y_k)\|(\|\nabla J(y_k)\| - \sqrt{C_k\Gamma_k})$. In addition, consider $C_k\Gamma_k \geq \|\nabla J(y_k) - \nabla\mu_k\|^2 \geq \|\mu_k\|^2 - \|\nabla J(y_k)\|^2$, we have $\beta\alpha^2\|\mu_k\|^2 \leq \beta\alpha^2(C_k\Gamma_k + \|\nabla J(y_k)\|^2)$. Particularly, we can rewrite $\|\nabla J(y_k)\|\|\Lambda_k\| + \beta\|\Lambda_k\|$ as $\beta(\|\Lambda_k\| - \frac{1}{2\beta}\|\nabla J(y_k)\|)^2 - \frac{1}{4\beta}\|\nabla J(y_k)\|^2$. Combining the above discussions, we have $J(y_{k+1}) - J(y_k) \leq -\alpha[(1 + \frac{1}{4\beta\alpha})\|\nabla J(y_k)\|^2 - (\sqrt{C_k\Gamma_k}\|\nabla J(y_k)\| + \beta\alpha(C_k\Gamma_k + \|\nabla J(y_k)\|^2) + \frac{\beta}{\alpha}(\|\Lambda_k\| + \frac{1}{2\beta}\|\nabla J(y_k)\|^2))]$. To make $J(y_{k+1}) - J(y_k) \leq 0$, let $\sqrt{C_k\Gamma_k} \leq \|\nabla J(y_k)\|/5$ and $\|\Lambda_k\| \leq \frac{\|\nabla J(y_k)\|}{10\beta}$. Then, by the definition of $\mathcal{Y}_k$, we have $\frac{m\theta_0^2}{\theta_1^2} - \frac{\rho_1^2\phi_{SE}(\rho_1)}{\theta_1^4(\theta_0^2\epsilon+\sigma_g^2/N_k)} \leq \frac{\|\nabla J(y_k)\|^2}{25C_k}$. In particular, when $\rho_1 \in (0, \theta_1\sqrt{W(\frac{\xi_k}{\theta_0^4\theta_1^2})})$, it can be derived that $\rho_1^2\phi_{SE}(\rho_1) \geq \xi_k$ where $\xi_k := (\theta_1^2(\theta_0^2\epsilon + \sigma_g/N_k)(25mC_k\theta_0^2 - \theta_1^2\|\nabla J(y_k)\|^2))/(25C_k)$. Since $\rho_1 \leq \rho_2 \leq \theta_1$, we can obtain that $\xi_k \leq \theta_0^4\theta_1^2$. Therefore, it can be derived that $\theta_0^2 \leq m\sigma_g/(N_k(1-m\epsilon))$ and $\theta_1^2 \geq \frac{25\theta_0^2C_k}{\|\nabla J(y_k)\|^2}(m - \frac{\theta_0^2}{\theta_0^2\epsilon+\sigma_g/N_k})$ when $0 \leq \epsilon \leq \frac{1}{m}$, and thus, the proof of Theorem 1 is complete. □

Remark 1. Note that the inequalities of the hyperparameters θ_0 and θ_1 in Theorem 1 will be applied in the optimization of the Jacobian Gaussian process hyperparameters. Typically, this is achieved by imposing constraints during the optimization of the marginal likelihood (Rasmussen & Williams, 2006).

3.2. ESO-based event-triggered tracking control

Considering the limited communication resources in network control systems, we develop a class of event-triggered controllers for the subsystems in the lower layer. In particular, the design of ETC in this work follows an emulation-based approach..

Firstly, to tackle the unknown disturbances and unmodeled dynamics of system (1), a linear ESO (Gao, 2003) is developed for each subsystem:

$$\dot{\hat{x}}_i = \hat{A}\hat{x}_i + \hat{D}(\varepsilon_i)\hat{H}_i\tilde{x}_{i1} + \hat{B}u_i,\ i = 1, \ldots, m, \tag{7}$$

where $\hat{x}_i = [\hat{x}_{i1}, \ldots, \hat{x}_{in}, \hat{x}_{i(n+1)}]^{\mathrm{T}}$ and $\hat{x}_{i(n+1)}$ is the observation of the extended state which is defined by $x_{i(n+1)} = f_i(t, x_i, w_i)$, $\hat{A} = [0\ I_n; 0\ 0]_{(n+1)\times(n+1)}$, $\hat{B} = [0, \ldots, 0, 1, 0]^{\mathrm{T}}_{(n+1)\times 1}$, $\hat{D}(\varepsilon_i) = \mathrm{diag}\{1/\varepsilon_i, 1/\varepsilon_i^2, \ldots, 1/\varepsilon_i^{n+1}\}$. Note that ε_i's are the gain parameters within $(0, 1)$, which are utilized to amplify the observer's gain, thereby suppressing uncertainties and enabling rapid convergence of the observation error. Additionally, to guarantee the exponential stability of the nominal observer (without disturbances and uncertainties), the observer gain $\hat{H}_i = [h_{i1}, \ldots, h_{in}, h_{i(n+1)}]^{\mathrm{T}}$ is generally designed to make the matrix $E_i = \hat{A} - \hat{H}_i\hat{C}$ Hurwitz, where $\hat{C} = [1, 0, \ldots, 0]_{1\times(n+1)}$.

Remark 2. Besides the high-gain ESO in (7), sliding mode control (SMC) and internal model control (IMC) are also effective for handling disturbances and uncertainties. SMC maintains system state on a designed sliding surface but imposes stringent requirements on hardware and may result in chattering phenomenon. IMC needs precise system models and state information, which makes it not well-suited for unknown systems. Meanwhile, high-gain ESO provides rapid disturbance compensation and is suitable for tracking control under uncertainties. Additionally, the peaking phenomenon in high-gain control can be alleviated by restricting the maximum control input amplitude.

Particularly, for $j = 1, \ldots, n+1$, let $\tilde{x}_{ij} := x_{ij} - \hat{x}_{ij}$, therewith we can define the scaled observation error $\eta_j = [\eta_{i1}, \ldots, \eta_{i(n+1)}]$ as $\eta_{ij} = \tilde{x}_{ij}/\varepsilon_i^{n+1-j}$. Hence, the dynamics of η_i can be derived as:

$$\begin{aligned}\dot{\eta}_{ij} &= [\eta_{i(j+1)} - h_{ij}\eta_{i1}]/\varepsilon_i, j = 1, \ldots, n\\ \dot{\eta}_{i(n+1)} &= \dot{x}_{i(n+1)} - h_{i(n+1)}\eta_{i1}/\varepsilon_i,\end{aligned} \tag{8}$$

In this work, the reference generated by the upper layer evolves in discrete time. To mitigate the impact of the abrupt changes in the reference, the introduction of the tracking differentiator (TD) proves to be an effective approach (Han, 2009; Han & Wang, 1994). Herein, we consider a TD for each subsystem as follows (Guo & Zhao, 2011a):

$$\dot{r}_i(t) = \tilde{A}_i r_i(t) + \tilde{B}_i \mathring{y}_i(t),\ \ i = 1, \ldots, m \tag{9}$$

where $r_i(t) = [r_{i1}, \ldots, r_{i(n+1)}]^{\mathrm{T}}$, $\tilde{A}_i = [0\ I_{n\times n}; -p_{i1}R_i^{n+1}, \ldots, -p_{i(n+1)}R_i]$ is Hurwitz with $R_i > 0$, $\tilde{B}_i = [0, \ldots, 0, p_{i1}R_i^{n+1}]^{\mathrm{T}}$. For the convergence of the tracking differentiator (9), we have the following lemma.

Lemma 2. *Considering that the signal $\mathring{y}_i(t)$ remains bounded and continuously differentiable in the time interval $(t_{i1}, t_{i2}]$, $t_{i1} < t_{i2} < \infty$. There exist a $\tilde{R}_i > 0$ and an R_i-dependent $t_{R_i} \in (t_{i1}, t_{i2}]$ such that for all $\infty > R_i \geq \tilde{R}_i$ and $t \in [t_{R_i}, t_{i2}]$, $|r_{i1}(t) - \mathring{y}_i(t)| \leq \delta_{R_i}$ where $\delta_{R_i} > 0$. In particular, it follows $|r_{ij}(t)| \leq \Gamma_i(R_i)$ with $\Gamma_i(R_i) : \mathbb{R} \to \mathbb{R}$.*

Remark 3. Lemma 2 indicates that for the step changes in the input signal of TD in (10), there exists a sufficiently large R_i such that the output of TD can approximate the input with a sufficiently small error in finite time. In general, the choice of R_i can be determined by considering the capabilities of the practical systems as well as the performance of the controllers. Note that Lemma 2 is an application of Theorem 3.2 in Guo and Zhao (2011a) when the input signal is piecewise constant and different from the initial value of TD. Hence, the proof of Lemma 2 can be completed following a similar line of arguments as the proof of Theorem 3.2 in Guo and Zhao (2011a).

Then, based on the ESO proposed in (7) and the TD in (9), the output-based feedback controller is designed as follows:

$$\bar{u}_i = -\textstyle\sum_{j=1}^{n} l_{ij}(\hat{x}_{ij} - r_{ij}) - \hat{x}_{i(n+1)} + r_{i(n+1)}, \tag{10}$$

where l_{ij}'s are the control gains. Define $L_i = [l_{i1}, \ldots, l_{in}]$, thus the control gains need to be designed such that $F_i = A - BL_i$ is Hurwitz. Particularly, a limitation of the control signal amplitude is considered:

$$\tilde{u}_i(t) = \tilde{M}_i \mathrm{sat}\left(\frac{\bar{u}_i(t_{ik})}{\tilde{M}_i}\right), \tag{11}$$

where $\mathrm{sat}(v) = \mathrm{sign}(v)\min\{1, |v|\}$ and $\tilde{M}_i$'s are positive constants which limits the maximum value of the control signals to suppress the peaking phenomenon arising in the high-gain approach (Freidovich & Khalil, 2008). For example, considering the conditions of Lemma 3, $\tilde{M}_i$ can be chosen to satisfy $\tilde{M}_i \leq \min_{e_i\in\mathcal{S}_{i0,e}, \eta_i\in\mathcal{S}_{i0,\eta}, |r_{ij}|\leq\Gamma(R_i)}(10)$.

Moreover, we consider the networked control scenario with event-triggered signal transmission strategy. For each subsystem, we consider the event-triggering mechanism in the sensor–observer channel (S/O) and controller–actuator (C/A) channel, which are designed in the following form:

$$t_{i(s+1),y} = \inf_{t>t_{is,y}} \left\{t : |\sigma_{i,y}(t)| \geq b_i\varepsilon_i\right\}, \tag{12}$$

$$t_{i(s+1),u} = \inf_{t>t_{is,u}} \left\{t : |\sigma_{i,u}(t)| \geq \frac{c_i\varepsilon_i}{(1+|\tilde{u}_i(t_{is,u})|)^{d_i}}\right\}, \tag{13}$$

where $t_{is,y}$, $t_{i(s+1),y}$ and $t_{is,u}$, $t_{i(s+1),u}$ represent the consecutive triggering instants for the ith subsystem in S/O channel and C/A channel, respectively; $\sigma_{i,y}(t) = (y_i(t) - y_i(t_{is,y}))/\varepsilon_i^n$ is the ETM-induced sampling error of the system outputs; $\sigma_{i,u}(t) = \tilde{u}_i(t) - \tilde{u}_i(t_{is,u})$ is the ETM-induced sampling error of the control input, and b_i, c_i, $d_i > 0$ are triggering threshold parameters.

Remark 4. The ETM in (12) is designed based on the principle of 'send-on-delta' as it determines the output transmission time by the difference between the current output and the last transmitted output, with the frequency of transmissions regulated by b_i. The ETM for the C/A channel is designed as a power function, aiming to meet rapid tracking control needs. This guarantees frequent updates when the controller output is large, thus maintaining tracking speed, and reduces update frequency as the system approaches the setpoint. Particularly, c_i controls the triggering threshold and d_i controls the decay rate. The design of c_i typically involves an emulation-based approach, while the selection of d_i is informed by empirical methods.

Considering the ETMs in (12) and (13), the ESO in (7) can be reformulated as:

$$\dot{\hat{x}}_i = \hat{A}\hat{x}_i + \hat{D}(\varepsilon_i)\hat{H}_i(\tilde{x}_{i1} - \sigma_{i,y}) + \hat{B}(u_i - \sigma_{i,u}). \tag{14}$$

Then, the dynamics of the observation error in (8) can be rewritten as:

$$\begin{aligned}\dot{\eta}_{ij} &= [\eta_{i(j+1)} - h_{ij}(\eta_{i1} + \sigma_{i,y})]/\varepsilon_i, j = 1, \ldots, n,\\ \dot{\eta}_{i(n+1)} &= \dot{x}_{i(n+1)} - h_{i(n+1)}(\eta_{i1} + \sigma_{i,y})/\varepsilon_i,\end{aligned} \tag{15}$$

In addition, we can define the tracking error as $e_{ij} = x_{ij} - r_{ij}$, $j = 1, \ldots, n$. Furthermore, we can obtain the dynamics of the tracking error according to (1) and (11) as

$$\dot{e}_i = Ae_i + B(x_{i(n+1)} + u_i - \sigma_{i,u} - r_{i(n+1)}), \tag{16}$$

where $e_i = [e_{i1}, e_{i2}, \ldots, e_{in}]^{\mathrm{T}}$.

Define two Lyapunov candidates as $V_1(\eta) = \sum_{i=1}^{m} V_{1i}(\eta_i) = \sum_{i=1}^{m} \eta_i^{\mathrm{T}} P_i \eta_i$ and $V_2(e) = \sum_{i=1}^{m} V_{2i}(e_i) = \sum_{i=1}^{m} e_i^{\mathrm{T}} Q_i e_i$, where P_i's and Q_i's are calculated by solving Lyapunov equations $P_i E_i + E_i^{\mathrm{T}} P_i = -I_{(n+1)\times(n+1)}$ and $Q_i F_i + F_i^{\mathrm{T}} Q_i = -I_{n\times n}$. Then, we can define the sets $\mathcal{I}_i := \{\eta_i \in \mathbb{R}^{n+1} : V_{1i}(\eta_i) \leq m_{1i}\varepsilon_i^2\}$ and $\mathcal{E}_i := \{e_i \in \mathbb{R}^n : V_{2i}(e_i) \leq m_{2i}\}$, where m_{1i}, m_{2i} are positive constants. Subsequently, the semi-global stability of the proposed event-triggered controller are provided in the following lemma.

Lemma 3. *Consider the system (1) satisfying Assumptions 1–2. Providing continuously differentiable and bounded references $r_i(t)$, then under the ETMs (12) and (13), the proposed event-triggered ESO (14) and output-based feedback controller (11) guarantee that there exist $\varepsilon_i^* > 0$ and $t_{\varepsilon_i} > 0$ such that $\eta_i(t) \in \mathcal{I}_i$, $e_i(t) \in \mathcal{E}_i$ for all $t \geq t_0 + t_{\varepsilon_i}$, $\varepsilon_i \in (0, \varepsilon_i^*)$, $\|\eta_i(t_0)\| \leq C_{\eta_i,0}$, and $\|e_i(t_0)\| \leq C_{e_i,0}$ with $C_{\eta_i,0}, C_{e_i,0} \in \mathbb{R}^+$.*

Proof. Please see the Appendix.

Remark 5. Lemma 3 shows that a sufficiently small ε_i can guarantee the convergence of the system's tracking error. According to Lemma 3, due to the presence of unknown nonlinearities, providing an analytical form for ε_i would be overly complex and challenging to verify directly. Therefore, a trial-and-error approach is typically used to determine the value of ε_i. Note that as ε_i decreases, more satisfactory control performance can be achieved, but this also leads to more severe peaking phenomenon and increases the frequency of transmission events. Thus, the selection of parameter ε_i is a trade-off between control performance and transmission frequency.

Remark 6. With the multiple parameters involved in the ETC controller proposed in this work, the co-design of these parameters presents a notable challenge. Therefore, we provide a feasible emulation-based parameter design process: (1) Design the observer and controller gains to satisfy the Hurwitz condition; (2) Select appropriate gain parameters ε_i to guarantee the closed-loop control performance; (3) Design the event-triggering parameters b_i, c_i, and d_i to reduce the number of triggering events to the extent possible.

Lemma 4. *For all the subsystems, under the same conditions of Lemma 3, the inter-event time of the ETMs (12) and (13) is lower bounded by $t_{i(s+1),y} - t_{is,y} \geq \Delta_{miet,y}$ and $t_{i(s+1),u} - t_{is,u} \geq \Delta_{miet,u}$, respectively.*

Proof. From Lemma 3, since e_i converge to a compact set $\mathcal{E}_i$, the states of each subsystem are bounded, i.e., $\|x_i(t)\| \leq C_{x_i}$ with $C_{x_i} > 0$. Then, we have $y_i(t) - y_i(t_{is,y}) = \int_{t_{is,y}}^{t} \dot{x}_{i1}(\tau)\mathrm{d}\tau = \int_{t_{is,y}}^{t} x_{i2}(\tau)\mathrm{d}\tau$. Therefore, by $|x_{i2}(t)| \leq \|x_i(t)\| \leq C_{x_i}$, we have $y_i(t) - y_i(t_{is,y}) \leq C_{x_i}(t - t_{is,y})$. According to (12), we have $C_{x_i}(t_{i(s+1),y} - t_{is,y}) \geq |y_{i(s+1),y} - y_{is,y}| \geq b_i\varepsilon_i^{n+1}$. Let $\Delta_{miet,y} := b_i\varepsilon_i^{n+1}/C_{x_i}$, then we have $t_{i(s+1),y} - t_{is,y} \geq \Delta_{miet,y} > 0$.

Similarly, by the definition of the $\sigma_{i,u}(t)$, we have $\sigma_i(t) \leq |\sigma_i(t)| \leq \int_{t_{ik}}^{t} \left|\dot{\tilde{u}}_i(\tau)\right| \mathrm{d}\tau$. In particular, by Lemma 3, we have $\left|\dot{\tilde{u}}_i(t)\right| \leq \sum_{j=1}^{n} k_{ij}(\varepsilon_i^{n+1-j}|\dot{\eta}_i| + |\dot{e}_i| + |\dot{\hat{x}}_{i(n+1)}| + |\dot{r}_{i(n+1)}|)$ for each subsystem. Since η_i and e_i are both bounded, there exist a positive constants $C_{u_i} > 0$ such that $|\dot{\tilde{u}}_i(t)| \leq C_{u_i}$. Taking $t = t_{i(s+1),u}$, by (13), we have $\frac{c_i\varepsilon_i}{(1+|\tilde{u}_i(t_{is,u})|)^{d_i}} \leq C_{u_i}(t_{i(s+1,u)} - t_{is,u})$. Hence, by (11), we can readily obtain that $t_{i(s+1),u} - t_{is,u} \geq \Delta_{miet,u} > 0$, where $\Delta_{miet,u} := c_i\varepsilon_i/C_{u_i}(1+\tilde{M}_i)^{d_i}$. □

Remark 7. Lemma 4 builds on Lemma 3 by leveraging the boundedness of the closed-loop system and the observer states to derive the positive minimum inter-event times for the two proposed event-triggering mechanisms (12) and (13), which are only affected by the system and observer states and do not converge to zero with time, thus excluding Zeno behavior.

Next, we provide the convergence of Algorithm 1 under the developed event-triggered control strategy in the following theorem.

Theorem 2. *Under the conditions of Theorem 1 and Lemma 3. For each subsystem, let the threshold parameter b_i in (12) satisfy $b_i > \sqrt{-2\theta_1 \ln(\epsilon)/\varepsilon_i^{2n+2}}$ for all $k \geq 1$. Then, there exist $\tilde{\varepsilon}_i^* > 0$ and $\tilde{t}_{\varepsilon_i} > 0$ such that for all $\varepsilon_i \in (0, \tilde{\varepsilon}_i^*)$ and $\delta \in (0,1)$, $\|\Lambda_k\| \leq \|\nabla J(y_k)\|/10\beta$ and $J(y_{k+1}) - J(y_k) \leq 0$ hold with probability at least $1-\delta$.*

Proof. Referring to Theorem 1, the data in the training set needs to satisfy $\min_{y^{\{i\}}, y^{\{j\}} \in \mathcal{D}_k, i\neq j} \|y^{\{i\}} - y^{\{j\}}\|^2 \geq -2\theta_1^2 \ln(\epsilon)$. In addition, we have $\|y^{\{i\}} - y^{\{j\}}\|^2 = \sum_{i=1}^{m}(y_i^{\{i\}} - y_i^{\{j\}})^2$. Then, by (12), we have $(y_i^{\{i\}} - y_i^{\{j\}})^2 \geq b_i^2\varepsilon_i^{2(n+1)}$. When $b_i > \sqrt{-2\theta_1 \ln(\epsilon)/\varepsilon_i^{2n+2}}$, it follows $\|y^{\{i\}} - y^{\{j\}}\|^2 \geq (y_i^{\{i\}} - y_i^{\{j\}})^2 \geq -2\theta_1 \ln(\epsilon)$. Hence the minimum distance condition of the training samples is satisfied.

Particularly, if $m_{2i} \leq \frac{\underline{\lambda}(Q_i)}{100m\beta^2}\|\nabla J(y_k)\|^2$, by Lemma 3, there exist $\tilde{\varepsilon}_i^* > 0$ and $\tilde{t}_{\varepsilon_i} > 0$ such that for all $\varepsilon_i \in (0, \tilde{\varepsilon}_i^*)$ and $t > t_{i0} + \tilde{t}_{\varepsilon_i}$, $e_i(t) \in \mathcal{E}_i$. Then, we have $\|e_i\|^2 \leq \frac{m_{2i}}{\underline{\lambda}(Q_i)} \leq \frac{\alpha^2}{m}(\|\nabla J(y_k)\|^2 - C_k\Gamma_k)$ holds with probability at least $1-\delta$. Hence, with the same probability, we can obtain that $\|\tilde{e}_k\| = \sqrt{\sum_{i=1}^{m}\|e_i\|^2} \leq \sqrt{m}\frac{\|\nabla J(y_k)\|}{10\beta}$. By Theorem 1, we have $J(y_{k+1}) - J(y_k) \leq 0$ holds with probability at least $1-\delta$, which completes the proof. □

Remark 8. Note that Theorem 2 provides the design condition for the triggering parameters b_i in the S/O channel, which ensures that the minimum distance condition of the training data in Theorem 1 can be reliably satisfied. This demonstrates the advantage of using event-triggered mechanisms in learning-based gradient estimation algorithms. Moreover, under the condition specified by Theorem 2, the selection of b_i is a trade-off between the control performance and transmission frequency following an emulation-based design approach.

Remark 9. Referring to Lemma 3, to guarantee the overall tracking error satisfies $\|\Lambda_k\| \leq \|\nabla J(y_k)\|/10\beta$, we can select $m_{2i} < \underline{\lambda}(Q_i)\|\nabla J(y_k)\|^2/(100m\beta^2)$. In particular, according to Algorithm 1, from a practical standpoint, $\|J(y_k)\|^2$ can be replaced with M_{grad}, thereby making it more feasible to determine ε_i accordingly.

4. Numerical example

In this section, we show the effectiveness of the developed Jacobian GP-based ESC method through numerical examples. Firstly, we consider a system with 4 subsystems: $\dot{x}_{i1} = x_{i2}$, $\dot{x}_{i2} = t^2e^{-t} + x_{i1} + x_{i2} + w_i + u_i$, $i = 1, \ldots, 4$, where $w_i = \cos(2t+1)$ is the external disturbances. The outputs are $y_1(t) = x_{11}(t)$, $y_2(t) = x_{21}(t)$, $y_3(t) = x_{31}(t)$, and $y_4(t) = x_{41}(t)$. Then, the cost function is defined as $J(y(t)) = (y_1(t)-6)^2 + (y_2(t)-2)^2 + (y_3(t)-2)^2 + (y_4(t)-1)^2$. In this example, the value of cost function $J(y(t))$ reaches its minimum 0 when $y_1 = 6$, $y_2 = 2$, $y_3 = 2$, and $y_4 = 1$. The output measurements of the cost are affected by Gaussian noises with zero means and standard deviation $\sigma_g = 0.01$. To compromise learning performance with computation complexity, the simulation step is set to be 0.01s.

For each subsystem, the observer gains are designed as $h_{i1} = 3$, $h_{i2} = 3$, and the gain parameter is $\varepsilon_i = 0.02$. The control gains

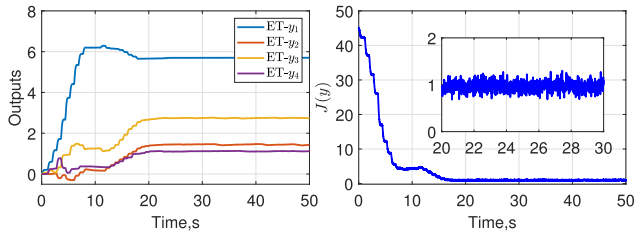


Fig. 2. The control performance and the optimization results of Algorithm 1 with $b_i = 500$, $c_i = 150$, $d_i = 1.2$.

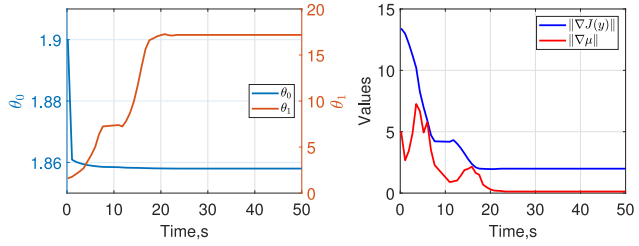


Fig. 3. The selected hyperparameters θ_0, θ_1 and the evolution of the Euclidean norm of the true gradient $\nabla J(y)$ and the estimated gradient $\nabla\mu$.

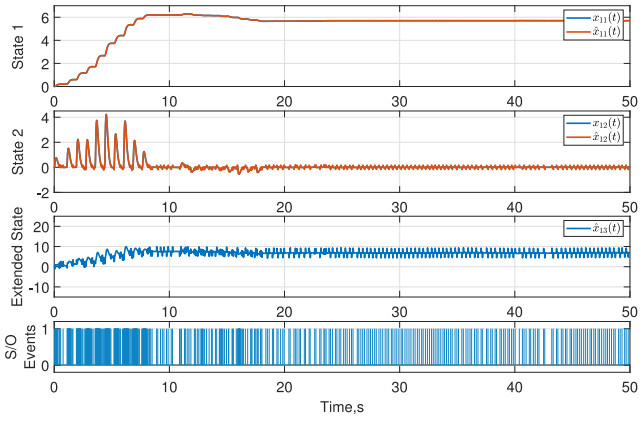


Fig. 4. The observation performance of the subsystem 1.

are designed as $l_{i1} = 5$, $l_{i2} = 4$. Fig. 2 illustrates the results of output tracking for each subsystem and the descent of the unknown cost function. Fig. 3 provides the hyperparameters utilized in the simulation along with the relationship between the estimated gradient and the true gradient. Particularly, the validation of the hyperparameters was completed by using the conditions in Theorem 1 as the constraints in maximizing the negative log marginal likelihood of the Jacobian GP. Fig. 4 demonstrates the observation performance of the developed event-triggered high-gain ESO by taking subsystem 1 as an example. It can be observed that, even with a limited number of events, the observer is still able to maintain satisfactory observation performance.

For comparison, a sliding mode controller (SMC) (Zhang & Ordóñez, 2009) is also incorporated into the developed framework (see Fig. 5). Additionally, we consider the impact of events on the optimization rate. By taking subsystem 1 as an example, we consider two ESO design with different parameter settings (namely, ESO1 and ESO2); specifically, the parameters of ESO1 are designed as $b_i = 500, c_i = 150, d_i = 1.2$, and the parameters of ESO2 are designed as $b_i = 2500, c_i = 40, d_i = 0.5$. Table 1 presents comparison results on the number of trigger events from the S/O and C/A channels, the convergence time (the time at which optimization stops), and the final extremum

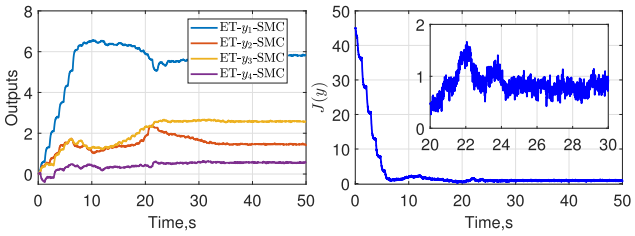


Fig. 5. The control performance and the optimization results of the developed framework under sliding-mode control (Zhang & Ordóñez, 2009) with $b_i = 2500$, $c_i = 900$, $d_i = 0.5$.

Table 1
Comparisons with a less events case as well as the SMC control strategy.

Methods	S/O Events	C/A Events	Convergence time	Extremum
ESO1	665	2024	23.28 s	0.8858
ESO2	425	3114	20.82 s	0.7873
SMC	546	2548	35.32 s	0.7194

achieved. It can be observed that the sliding mode control utilizes more events to achieve rapid convergence, which stems from its characteristic of rapid switching on the sliding mode surface. On the other hand, for the developed high-gain ESO method, the number of transmissions in the S/O channel is decreased by approximately 86.7%, and in the C/A channel by approximately 59.5%. However, the speed of optimization convergence decreases (from 20.82 s to 23.28 s) when the number of events is reduced. Thus, the results indicate that the selection of event-triggering parameters is a trade-off between transmission performance and convergence rate.

5. Conclusion

This paper proposes an event-triggered extremum seeking control framework based on Jacobian GP learning. We ensure the high-probability convergence of the framework by deriving hyperparameters optimization conditions, devising strategies for selecting training samples, and designing the event-triggered controller for stability. As demonstrated in the numerical examples, the method developed in this paper can achieve model-free optimization of unknown performance functions under event-triggered transmission strategy. In future, our research will explore more general forms of kernel functions to broaden the application scope of this framework, while also taking into account more complex disturbances such as unmatched or stochastic ones.

Appendix. Proof of Lemma 3

Initially, by calculating the derivatives of V_{i1} and V_{i2}, we have $\dot{V}_{1i} = \sum_{j=1}^{n} \frac{\partial V_{1i}}{\partial \eta_{ij}}[\eta_{i(j+1)} - h_{ij}\eta_{i1}]/\varepsilon_i + \frac{\partial V_{1i}}{\partial \eta_{i(n+1)}}[\dot{x}_{i(n+1)} - h_{i(n+1)}\eta_{i1}/\varepsilon_i] + \sum_{j=1}^{n+1} \frac{\partial V_{1i}}{\partial \eta_{ij}} h_{ij}\sigma_{i,y}$ and $\dot{V}_{2i} = \sum_{j=1}^{n-1} \frac{\partial V_{2i}}{\partial e_{ij}} e_{i(j+1)} + \frac{\partial V_{2i}}{\partial e_{in}}[x_{i(n+1)} + \tilde{u}_i - \sigma_i - r_{i(n+1)}]$. Then, we can readily obtain that $\underline{\lambda}(P_i)\|\eta_i\|^2 \leq V_{1i}(\eta_i) \leq \overline{\lambda}(P_i)\|\eta_i\|^2$, $|\partial V_{1i}(\eta_i)/\partial \eta_i| \leq 2\overline{\lambda}(P_i)\|\eta_i\|$, $\underline{\lambda}(Q_i)\|e_i\|^2 \leq V_{2i}(e) \leq \overline{\lambda}(Q_i)\|e_i\|^2$, $|\partial V_{2i}(e_i)/\partial e_i| \leq 2\overline{\lambda}(Q_i)\|e_i\|$. Based on the initial condition of $\eta_i(t_0)$ and $e_i(t_0)$, we can define two compact sets $\mathcal{S}_{i0,\eta} := \left\{\eta_i \in \mathbb{R}^{n+1} | V_{i1}(\eta_i) \leq \underline{\lambda}(P_i) C_{\eta_i,0}^2\right\}$ and $\mathcal{S}_{i0,e} := \left\{e_i \in \mathbb{R}^n | V_{i2}(e_i) \leq \underline{\lambda}(Q_i) C_{e_i,0}^2\right\}$. Thus, we have $\eta_i(t_0) \in \mathcal{S}_{i0,\eta}$ and $e_i(t_0) \in \mathcal{S}_{i0,e}$. Then, the proof of Lemma 3 will proceed from discussions of the following circumstances: (i) $\eta_i \in \partial\mathcal{I}_i$, $e_i \in \mathcal{S}_{i0,e}$; (ii) $\eta_i \notin \mathcal{I}_i$, $e_i \in \text{int}(\mathcal{S}_{i0,e})$; (iii) $\eta_i \in \mathcal{I}_i$, $e_i \in \partial\mathcal{S}_{i0,e}$; (iv) $\eta_i \in \mathcal{I}_i$, $e_i \notin \mathcal{E}_i$.

Step 1: When $\eta_i \in \partial\mathcal{I}_i$, $e_i \in \mathcal{S}_{i0,e}$, the derivatives of the extended state $\dot{x}_{i(n+1)}$ can be calculated as $\dot{x}_{i(n+1)} = \frac{\partial f_i}{\partial t} + \sum_{j=1}^{n-1} \frac{\partial f_i}{\partial x_{ij}} x_{i(j+1)} + \frac{\partial f_i}{\partial x_{in}}[f_i + \tilde{u}_i - \sigma_i] + \frac{\partial f_i}{\partial w_i}\dot{w}_i$. According to Lemma 1, for all $j \leq n-1$, we have $x_{i(j+1)} \leq \|x_i\| \leq \|e_i\| + \mathcal{F}_i(R_i)$. Then, according to Assumptions 1–2 and (11), we have $\|\kappa_i(x_i, w_i)\| \leq c_f$ with $c_f > 0$. In addition, by (13), we can obtain that $|\sigma_i(t)| \leq c_i\varepsilon_i$. Let $\tilde{K}_i := \sum_{j=1}^{n} k_{ij}$, we can obtain that $f_i + \tilde{u}_i - \sigma_i \leq \tilde{K}_i(\varepsilon_i\|\eta_i\| + \|e_i\|) + \|\eta_i\| + \mathcal{F}_i(R_i) + c_i\varepsilon_i$. Hence, we can conclude that the derivative of the extended state satisfies $\dot{x}_{i(n+1)} \leq c_f\|\eta_i\| + c_f\tilde{K}_i\varepsilon_i\|\eta_i\| + D_1\|e_i\| + D_2 + c_f n\mathcal{F}_i(R_i) + c_f c_i\varepsilon_i$ where $D_1 = c_f(n - 1 + \tilde{K}_i)$ and $D_2 = c_f(1 + c_w)$]. Subsequently, according to (8), we can obtain that $\sum_{j=1}^{n} \frac{\partial V_{1i}}{\partial \eta_{ij}}(\eta_{i(j+1)} - h_{ij}\eta_{i1}) - \frac{\partial V_{1i}}{\partial \eta_{i(n+1)}} h_{i(n+1)}\eta_{i1} = 2\eta_i^{\mathrm{T}} P_i(\hat{A} + \hat{B}\hat{H})\eta_i \leq -\|\eta_i\|^2$. In addition, from (12), we have $|\sigma_{i,y}| \leq b_i\varepsilon_i$. Thus, denoting $\varkappa_1 := 2\bar{\lambda}(P_i)$ and $\tilde{h}_i := \sum_{j=1}^{n+1} h_{ij}$, we have $\dot{V}_{1i} \leq -\frac{1}{\varepsilon_i}\|\eta_i\|^2 + \varkappa_1 D_1\|\eta_i\|\|e_i\| + \varkappa_1 c_f(1 + \tilde{K}_i\varepsilon_i)\|\eta_i\|^2 + \varkappa_1(D_2 + c_f n\mathcal{F}_i(R_i) + c_f c_i\varepsilon_i + \tilde{h}_i b_i)\|\eta_i\|$. By Young's inequality, we can derive that $\dot{V}_{1i} \leq -\frac{1}{\varepsilon_i}(1 - \varkappa_1 c_f(1 + \tilde{K}_i\varepsilon_i)\varepsilon_i - 3\varepsilon_i^{\frac{1}{3}} - \varepsilon_i^{\frac{4}{3}})\|\eta_i\|^2 + \varkappa_1^2\varepsilon_i^{\frac{2}{3}}\Big(D_1^2 C_{e_i 0} + D_2^2 + c_f^2 n^2 \mathcal{F}_i^2(R_i) + c_f^2 c_i^2 + \tilde{h}_i^2 b_i^2\Big)$. Subsequently, there exists an ε_{i1}^* such that $\dot{V}_{1i} \leq 0$.

Step 2: When $\eta_i \notin \mathcal{I}_i$ and $e_i \in \mathrm{int}(\mathcal{S}_{i0,e})$. Assume that $e_i \in \mathcal{S}_{i0,e}$ during the time interval $t \in [t_{i0}, t_{i1}]$. Referring to the fact that $-\|\eta_i\|^2 \leq -\frac{1}{\bar{\lambda}(P_i)} V_{1i}(\eta_i)$, we can obtain that $\dot{V}_{1i} \leq -\frac{\gamma_{i1}(\varepsilon_i)}{\bar{\lambda}(P_i)} V_{1i} + \gamma_{i2}(\varepsilon_i)$ where $\gamma_{i1}(\varepsilon_i) = -\frac{1}{\varepsilon_i}(1 - \varkappa_1 c_f(1 + \tilde{K}_i\varepsilon_i)\varepsilon_i - 3\varepsilon_i^{\frac{1}{3}} - \varepsilon_i^{\frac{4}{3}})$ and $\gamma_{i2}(\varepsilon_i) = \varkappa_1^2\varepsilon_i^{\frac{2}{3}}\Big(D_1^2 C_{e_i 0} + D_2^2 + c_f^2 n^2 \mathcal{F}_i^2(R_i) + c_f^2 c_i^2 + \tilde{h}_i^2 b_i^2\Big)$. Denoting $\delta t = t - t_{i0}$, we have $V_{1i}(\eta_i(t)) \leq \exp\Big(-\frac{\gamma_{i1}(\varepsilon_i)}{\bar{\lambda}(P_i)}\delta t\Big)\, V_{i1}(\eta_i(0)) + \frac{\gamma_{i2}(\varepsilon_i)}{\gamma_{i1}(\varepsilon_i)}\bar{\lambda}(P_i)\Big[1 - \exp\Big(-\frac{\gamma_{i1}(\varepsilon_i)}{\bar{\lambda}(P_i)}\delta t\Big)\Big]$. To guarantee that the tracking error e_i will stay in $\mathcal{S}_{i0,e}$, we can find an ε_{i2}^* and an ε_i-dependent $t_{\varepsilon_i,1}$ as $t_{\varepsilon_i,1} = -\frac{\bar{\lambda}(P_i)}{\gamma_{i1}(\varepsilon_i)} \ln\left(\frac{\tilde{m}_{1i}\varepsilon_i^2 - \frac{\gamma_{i2}(\varepsilon_i)}{\gamma_{i1}(\varepsilon_i)}\bar{\lambda}(P_i)}{\underline{\lambda}(P_i)C_{\eta_i,0}^2 - \frac{\gamma_{i2}(\varepsilon_i)}{\gamma_{i1}(\varepsilon_i)}\bar{\lambda}(P_i)}\right)$ such that for all $\varepsilon_i \in (0, \varepsilon_{i2}^*)$, $t_{\varepsilon_i,1} \geq 0$. Hence, the trajectory η_i will stay in $\mathcal{I}_i$ when $t > t_{i0} + t_{\varepsilon_i,1}$.

Step 3: Consider the scenario when $\eta_i \in \mathcal{I}_i$ and $e_i \in \partial\mathcal{S}_{i0,e}$. According to (10), we have $\sum_{j=1}^{n-1} \frac{\partial V_{2i}}{\partial e_{ij}} e_{i(j+1)} + \frac{\partial V_{2i}}{\partial e_{in}} \sum_{j=1}^{n} k_{ij} e_{ij} = 2e_i^{\mathrm{T}} Q(A + BK)e_i \leq -\|e_i\|^2$. In addition, for all $\eta_i \in \mathcal{I}_i$, it follows $V_{1i}(\eta_i) \leq \tilde{m}_{1i}\varepsilon_i^2$ and $\|\eta_i\| \leq \sqrt{\frac{\tilde{m}_{1i}}{\underline{\lambda}(P_i)}}\varepsilon_i$. Denoting $\tilde{K}_i = \sum_{j=1}^{n} k_{ij}$ and $\varkappa_2$, by Young's inequality, we conclude that $\dot{V}_{2i} \leq -\|e_i\|^2 + \varkappa_2(1 + \tilde{K}_i\varepsilon_i)\|e_i\|\|\eta_i\| + \varkappa_2 c_i\|e_i\| \leq -[1 - \varkappa_2^2(\tilde{K}_i^2\varepsilon_i^2 + 2)\varepsilon_i]\|e_i\|^2 + \Big(\frac{2m_{1i}}{\underline{\lambda}(P_i)} + c_i^2\Big)\varepsilon_i \leq -(1 - \gamma_{i3}(\varepsilon_i))\|e_i\|^2 + N_1\varepsilon_i$ where $\gamma_{i3}(\varepsilon_i) = \varkappa_2^2(\tilde{K}_i^2\varepsilon_i^2 + 2)\varepsilon_i$ and $N_1 := 2m_{1i}/\underline{\lambda}(P_i) + c_i^2$. Clearly, $\gamma_{i3}(\varepsilon_i) \to 0$ and $\gamma_{i4}(\varepsilon_i) \to 0$ as $\varepsilon_i \to 0$. Hence, there exists an ε_{i3}^* such that for all $\varepsilon_i \in (0, \varepsilon_{i3}^*)$, $\dot{V}_{i2} \leq 0$.

Step 4: When $\eta_i \in \mathcal{I}_i$ while $e_i \notin \mathcal{E}_i$. Similar to the line of the proofs in Step 2, we have $\dot{V}_{2i} \leq -\frac{1 - \gamma_{i3}(\varepsilon_i)}{\bar{\lambda}(Q_i)} V_{2i} + N_1\varepsilon_i$. Then, we can obtain that $V_{2i}(t) \leq \exp(-\frac{1 - \gamma_{i3}(\varepsilon_i)}{\bar{\lambda}(Q_i)}\delta t) V_{2i}(e_i(0)) + \frac{N_1\varepsilon_i}{1 - \gamma_{i3}(\varepsilon_i)}\bar{\lambda}(Q_i)(1 - \exp(-\frac{1 - \gamma_{i3}(\varepsilon_i)}{\bar{\lambda}(Q_i)}\delta t))$. From the characteristics of $\gamma_{i3}(\varepsilon_i)$, it is straightforward to conclude that there exist an ε_{i4}^* and an ε_i-dependent time $t_{\varepsilon_i,2} > 0$ such that for all $\varepsilon_i \in (0, \varepsilon_{i4}^*)$ and $t > t_{i0} + \tilde{t}_{\varepsilon_i,2}$, the trajectory of the tracking error will enter the set $\mathcal{E}_i$. Therefore, by defining $\varepsilon_i^* = \min\{\varepsilon_{i1}^*, \varepsilon_{i2}^*, \varepsilon_{i3}^*, \varepsilon_{i4}^*\}$ and $t_{\varepsilon_i} = \max\{t_{\varepsilon_i,1}, t_{\varepsilon_i,2}\}$, the proof of Lemma 3 is completed. □

References

Abdelrahim, M., Postoyan, R., Daafouz, J., & Nešić, D. (2017). Robust event-triggered output feedback controllers for nonlinear systems. *Automatica*, *75*, 96–108.

Beckers, T., Kulić, D., & Hirche, S. (2019). Stable Gaussian process based tracking control of Euler–Lagrange systems. *Automatica*, *103*, 390–397.

Freidovich, L. B., & Khalil, H. K. (2008). Performance recovery of feedback-linearization-based designs. *IEEE Transactions on Automatic Control*, *53*(10), 2324–2334.

Fu, L., & Özgüner, Ü. (2011). Extremum seeking with sliding mode gradient estimation and asymptotic regulation for a class of nonlinear systems. *Automatica*, *47*(12), 2595–2603.

Gao, Z. (2003). Scaling and bandwidth-parameterization based controller tuning. In *Proceedings of the 2003 American control conference, 2003: vol. 6*, (pp. 4989–4996). Denver, CO, USA: IEEE.

Girard, A. (2015). Dynamic triggering mechanisms for event-triggered control. *IEEE Transactions on Automatic Control*, *60*(7), 1992–1997.

Guay, M., Dochain, D., & Perrier, M. (2004). Adaptive extremum seeking control of continuous stirred tank bioreactors with unknown growth kinetics. *Automatica*, *40*(5), 881–888.

Guo, B.-Z., & Zhao, Z.-L. (2011a). On convergence of tracking differentiator. *International Journal of Control*, *84*(4), 693–701.

Guo, B.-Z., & Zhao, Z.-L. (2011b). On the convergence of an extended state observer for nonlinear systems with uncertainty. *Systems & Control Letters*, *60*(6), 420–430.

Han, J. (2009). From PID to active disturbance rejection control. *IEEE Transactions on Industrial Electronics*, *56*(3), 900–906.

Han, J., & Wang, W. (1994). Nonlinear tracking-differentiator. *Journal of System Science and Mathematical Science Chinese Series*, *14*(2), 177–183.

Heemels, W., & Donkers, M. (2013). Model-based periodic event-triggered control for linear systems. *Automatica*, *49*(3), 698–711.

Heemels, W. P. M. H., Johansson, K. H., & Tabuada, P. (2012). An introduction to event-triggered and self-triggered control. In *2012 IEEE 51st IEEE conference on decision and control* (pp. 3270–3285).

Heemels, W. P. M. H., Teel, A. R., van de Wouw, N., & Nešić, D. (2010). Networked control systems with communication constraints: tradeoffs between transmission intervals, delays and performance. *IEEE Transactions on Automatic Control*, *55*(8), 1781–1796.

Huang, J., Wang, W., Wen, C., & Li, G. (2020). Adaptive event-triggered control of nonlinear systems with controller and parameter estimator triggering. *IEEE Transactions on Automatic Control*, *65*(1), 318–324.

Krishnamoorthy, D., & Skogestad, S. (2022). Real-Time optimization as a feedback control problem – A review. *Computers & Chemical Engineering*, *161*, Article 107723.

Krstić, M., & Wang, H.-H. (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, *36*(4), 595–601.

Léchappé, V., Moulay, E., Plestan, F., & Han, Q.-L. (2019). Discrete predictor-based event-triggered control of networked control systems. *Automatica*, *107*, 281–288.

Li, F., & Liu, Y. (2020). Adaptive event-triggered output-feedback controller for uncertain nonlinear systems. *Automatica*, *117*, Article 109006.

Li, H., & Shi, Y. (2014). Event-triggered robust model predictive control of continuous-time nonlinear systems. *Automatica*, *50*(5), 1507–1513.

Li, T., Wen, C., Yang, J., Li, S., & Guo, L. (2020). Event-triggered tracking control for nonlinear systems subject to time-varying external disturbances. *Automatica*, *119*, Article 109070.

Liu, T., & Jiang, Z.-P. (2015). A small-gain approach to robust event-triggered control of nonlinear systems. *IEEE Transactions on Automatic Control*, *60*(8), 2072–2085.

Müller, S., von Rohr, A., & Trimpe, S. (2021). Local policy search with Bayesian optimization. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, & J. W. Vaughan (Eds.), *Advances in neural information processing systems: vol. 34*, (pp. 20708–20720). Curran Associates, Inc..

Nesic, D., & Teel, A. (2004). Input-output stability properties of networked control systems. *IEEE Transactions on Automatic Control*, *49*(10), 1650–1667.

Ran, M., Li, J., & Xie, L. (2021). A new extended state observer for uncertain nonlinear systems. *Automatica*, *131*, Article 109772.

Rasmussen, C. E., & Williams, C. K. I. (2006). Gaussian processes for machine learning. In *Adaptive computation and machine learning*, Cambridge, Mass: MIT Press.

Rodrigues, V. H. P., Oliveira, T. R., Hsu, L., Diagne, M., & Krstic, M. (2023). Event-Triggered extremum seeking control systems.

Srinivas, N., Krause, A., Kakade, S. M., & Seeger, M. W. (2012). Information-theoretic regret bounds for Gaussian Process optimization in the bandit setting. *Institute of Electrical and Electronics Engineers. Transactions on Information Theory*, *58*(5), 3250–3265.

Sun, H., Sun, J., & Chen, J. (2021). Analysis and synthesis of networked control systems with random network-induced delays and sampling intervals. *Automatica*, *125*, Article 109385.

Sun, J., Yang, J., & Zeng, Z. (2022). Predictor-based periodic event-triggered control for nonlinear uncertain systems with input delay. *Automatica*, *136*, Article 110055.

Suttner, R., & Krstić, M. (2023). Extremum seeking control for fully actuated mechanical systems on Lie groups in the absence of dissipation. *Automatica*, *152*, Article 110945.

Walsh, A., Forbes, J. R., Chee, S. A., & Ryan, J. J. (2017). Kalman-filter-based unconstrained and constrained extremum-seeking guidance on SO(3). *Journal of Guidance, Control, and Dynamics*, *40*(9), 2260–2271.

Wu, K., Kim, K., Garnett, R., & Gardner, J. (2023). The behavior and convergence of local Bayesian optimization. In *The 37th annual conference on neural information processing systems*. New Orleans, LA, USA.

Xie, X., Liu, Y., & Li, Q. (2023). Neural network-based adaptive event-triggered control for cyber–physical systems under resource constraints and hybrid cyberattacks. *Automatica*, *152*, Article 110977.

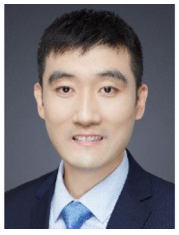
Ye, M., & Hu, G. (2016). A robust extremum seeking scheme for dynamic systems with uncertainties and disturbances. *Automatica*, *66*, 172–178.

Zhang, C., & Ordóñez, R. (2007). Numerical optimization-based extremum seeking control with application to ABS design. *IEEE Transactions on Automatic Control*, *52*(3), 454–467.

Zhang, C., & Ordóñez, R. (2009). Robust and adaptive design of numerical optimization-based extremum seeking control. *Automatica*, *45*(3), 634–646.

Zhu, Y., Fridman, E., & Oliveira, T. R. (2023). Sampled-data extremum seeking with constant delay: A time-delay approach. *IEEE Transactions on Automatic Control*, *68*(1), 432–439.

Jiliang Song was born in Shandong, China. He received the B.Eng. degree in automation from Chang'an University, Xi'an, China, in 2018, and the Ph.D. degree in control science and engineering from Beijing Institute of Technology, Beijing, China, in 2024. Currently, he is a Postdoctoral Fellow at the Beijing Institute of Control Engineering, Beijing, China. His research interests include event-based control, networked control systems, and model predictive control.



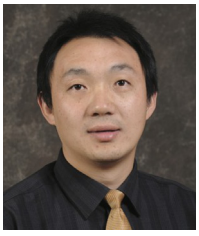
Dawei Shi is presently a Professor in the School of Automation at the Beijing Institute of Technology, Beijing, China. He received the B.Eng. degree in Electrical Engineering and Automation from the Beijing Institute of Technology in 2008, and the Ph.D. degree in control systems from the University of Alberta in 2014. In 2017–2018, he was a Postdoctoral Fellow in Bioengineering at the Harvard University.

His research focuses on the analysis and design of advanced sampled-data control systems, with applications to biomedical engineering, robotics, and motion systems. He serves as an Associate Editor/Technical Editor for several international journals, including Control Engineering Practice, IEEE/ASME Transactions on Mechatronics, IEEE Transactions on Industrial Electronics and IEEE Control Systems Letters.

Shu-Xia Tang received the Ph.D. degree in mechanical engineering from the Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA, USA, in 2016.

She is currently an Assistant Professor with the Department of Mechanical Engineering, Texas Tech University, Lubbock, TX, USA. Her research interests include stability analysis, estimation, and control design of distributed parameter systems.

Hao Yu received the B.S. and Ph.D. degrees in control theory and engineering from Beihang University, Beijing, China, in 2013 and 2018, respectively. From 2019 to 2022, he was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada. He is currently a Professor with the School of Automation, Beijing Institute of Technology, Beijing. His research interests include hybrid systems, event-triggered control, multiagent systems, data–driven PID control, adaptive control, and cyber–physical systems.



Yang Shi received the Ph.D. degree in electrical and computer engineering from the University of Alberta, Edmonton, AB, Canada, in 2005. From 2005 to 2009, he was an Assistant Professor and an Associate Professor in the Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, SK, Canada. In 2009, he joined the University of Victoria, Victoria, BC, Canada, where he is currently a Professor in the Department of Mechanical Engineering. His current research interests include networked and distributed systems, model predictive control (MPC), cyber–physical systems (CPS), robotics and mechatronics, navigation and control of autonomous systems (AUV and UAV), and energy system applications.

Dr. Shi received the University of Saskatchewan Student Union Teaching Excellence Award in 2007, and the Faculty of Engineering Teaching Excellence Award in 2012 at the University of Victoria (UVic). He is the recipient of the JSPS Invitation Fellowship (short-term) in 2013, the UVic Craigdarroch Silver Medal for Excellence in Research in 2015, the 2016 IEEE Transactions on Fuzzy Systems Outstanding Paper Award, the Humboldt Research Fellowship for Experienced Researchers in 2018. He is Vice-President on Conference Activities of IEEE IES, and Chair of IES Technical Committee of Industrial Cyber–Physical Systems. Currently, he is the Editor-in-Chief for IEEE Transactions on Industrial Electronics; he also serves as Associate Editor for Automatica, IEEE Transactions on Automatic Control, IEEE Transactions on Cybernetics, etc. He is General Chair of the 2019 International Symposium on Industrial Electronics (ISIE) and the 2021 International Conference on Industrial Cyber–Physical Systems (ICPS). He is a Fellow of IEEE, ASME, Engineering Institute of Canada (EIC), and Canadian Society for Mechanical Engineering (CSME), and a registered Professional Engineer in British Columbia, Canada.