This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3140259, IEEE Transactions on Automatic Control

Sampled-data Extremum Seeking with Constant Delay: A Time-delay Approach

Yang Zhu, Emilia Fridman, Tiago Roux Oliveira

***Abstract*—This paper proposes a constructive method for sampled-data extremum seeking (ES) with square wave dithers and constant delays, by using two time-delay approaches: one to averaging and the other to sampled-data control. We consider gradient-based ES for static maps which are of quadratic forms. By transforming the ES system to the time-delay system, we have developed a stability analysis via a Lyapunov-Krasovskii method. We derive the practical stability conditions in terms of linear matrix inequalities (LMIs) for the resulting time-delay system. The time-delay approach offers a quantitative calculation on the upper bound of the dither and sampling periods, constant delays that the ES system is able to tolerate, as well as the ultimate bound of the extremum seeking error. This is in the presence of uncertainties of extremum value and extremum point.**

I. INTRODUCTION

Extremum Seeking (ES) has proven itself as an effective online optimization technique due to its model-free feature. In 2000, Krstic and Wang introduced the first rigorous stability assessment of ES feedback in their publication [9], in which the averaging-based theory was employed, becoming a dominant tool for ES analysis in the literature. Since the 21 century, ES has seen its rapid growth both in terms of the theoretical developments and applications: ES via Lie bracket approximation [2], ES for PDE dynamics [12] and time-delay systems [11], ES under deception [6], resilient cooperative source seeking of multi-robot [5].

In parallel with continuous-time ES, sampled-data ES with discontinuous dithers is more friendly implementable in practice [17], [18]. The existing analysis approach to sampled-data ES models ES control systems as discrete-time systems [7], [10], [15] or hybrid systems [13], [14], [19]. Delays, that may stem from measurement or network communication, are unavoidable in practical applications. The convergence analyses based on the trajectory of discrete-time sequence or hybrid dynamical systems do not take into account delays. The existing theories of sampled-data ES cannot provide a quantitative analysis about how large the delays and the sampling intervals that ES control systems are able to withstand.

This work was supported by National Key R&D Program of China (Grant No. 2021YFB1714100, 2018YFA0703800), Science Fund for Creative Research Group of the National Natural Science Foundation of China (Grant No. 61621002), Zhejiang Key R&D Program (Grant No. 2021C01198), Zhejiang Jianbing Program (Grant No. 2022C01035), Post-Doctoral Program of the Planning and Budgeting Committee (PBC), Council for Higher Education in Israel, and Israel Science Foundation (grant no. 673/19)

Yang Zhu is with College of Control Science and Engineering, State Key Laboratory of Industrial Control Technology, Zhejiang University, China, and School of Electrical Engineering, Tel-Aviv University, Israel. zhuyang88@zju.edu.cn

Emilia Fridman is with School of Electrical Engineering, Tel-Aviv University, Israel. emilia@tauex.tau.ac.il

Tiago Roux Oliveira is with Department of Electronics and Telecommunication Engineering, State University of Rio de Janeiro, Rio de Janeiro, RJ 20550-900, Brazil. tiagoroux@uerj.br

The recent publication [4] presents a novel constructive approach to periodic averaging with efficient and quantitative bounds on the small parameter that preserves the stability of the original system provided the averaged system is stable.

Inspired by [4], an original time-delay approach for continuous-time ES with continuous dither of sine wave was proposed in [20]. In the present paper, we further expand the time-delay approach to sampled-data ES with discontinuous dither of square wave and constant delay which is important for implementation. Concentrating on the gradient-based ES of static quadratic maps, we transform the ES dynamics into a model with time-delay. The stability of the original ES plant is concluded from the stability of the resulting time-delay system. To find sufficient practical stability conditions in the form of LMIs, we construct a novel LKF which captures the full states of the closed-loop system. This improves in the examples the results of [20] that treated part of terms in the time-delay model as disturbances. Through the solution of the constructed LMIs, we find upper bounds on the dither period that ensures the practical stability, as well as the maximum sampling interval and delay that the ES control system is able to tolerate.

Different from the conventional averaging method [1], [16] and Lie bracket method [2], which are of "approximate" in essence, the time-delay method gives a precise conversion of the original ES system without any approximation. The ultimate bound on the extremum seeking error is also given in an explicitly quantitative way. It is important to see the impact of delays and sampling on the performance of ES algorithms. Moreover, our method suggests details for the selection of designer-specified parameters.

II. A Time-delay Approach to Sampled-data ES

We study sampled-data delayed implementation of gradient-based ES. We consider the periodic sampling with the constant delay. To avoid notational complexity, we address the case of two-input static map. The method can be extended to any $n > 2$ inputs by using the same arguments, but derivations are longer. For the results on one-input static map with square wave dithers, both in the continuous and the delayed sampled-data cases, see the companion conference paper [21].

A. ES System and a Time-delay Approach to Sampled-data

Consider the two-input static map $Q(\theta)$ given by

$$y(t) = Q(\theta(t)) = Q^* + \tfrac{1}{2}\big(\theta(t) - \theta^*\big)^T H\big(\theta(t) - \theta^*\big), \qquad (1)$$

where $y(t) \in \mathbb{R}$, $Q^* \in \mathbb{R}$, and

$$\theta(t) = [\theta_1(t), \theta_2(t)]^T, \quad \theta^* = [\theta_1^*, \theta_2^*]^T, \quad H = \begin{bmatrix} h_{11} & h_{12} \\ * & h_{22} \end{bmatrix}. \qquad (2)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3140259, IEEE Transactions on Automatic Control

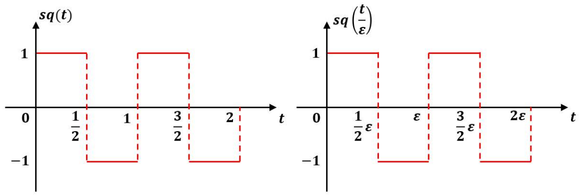


Fig. 1. The dither signal of the square wave

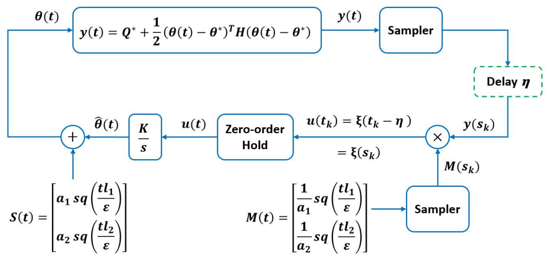


Fig. 2. The two-input sampled-data ES with discontinuous dither vectors $S(t)$ and $M(t)$ and constant delay η, where $u(t)=u(t_k)=\xi(s_k)=M(s_k)y(s_k)$ for $t\in[t_k,t_{k+1})$.

It is seen that the output of the quadratic map has a maximum (if $H<0$) or minimum (if $H>0$) value $y(t)=Q^*$ if the vector input $\theta(t)=\theta^*$. We define the real-time estimate $\hat{\theta}(t)$ of θ^* with the estimation error

$$\tilde{\theta}(t)=\hat{\theta}(t)-\theta^*. \tag{3}$$

The target of ES is to guarantee that $\hat{\theta}(t)$ converges towards θ^* so that the output $y(t)$ will converge to its optimal point Q^*. Usually, Q^*, H, θ^* are unknown, whereas the sign of the Hessian H is available. In the present paper, we assume the extremum point θ^* to be sought is uncertain from a known ball where each of its elements satisfies $\theta_i^*\in\left[\underline{\theta}_i^*,\overline{\theta}_i^*\right], i=1,2$ with $\sum_{i=1}^2\left(\overline{\theta}_i^*-\underline{\theta}_i^*\right)^2=\sigma_0^2$. The extremum value Q^* and the Hessian H are supposed to be known to derive efficient LMI conditions (as explained below, Q^* can be also uncertain).

As shown in Fig. 1, we introduce the square wave signal $\mathrm{sq}(t)$ which is of the form

$$\mathrm{sq}(t)=\begin{cases}1, & t\in\left[n,n+\frac{1}{2}\right),\\ -1, & t\in\left[n+\frac{1}{2},n+1\right),\end{cases}\quad n\in\mathbb{Z}_0^+. \tag{4}$$

As revealed in Fig. 2, the gradient-based sampled-data ES is given by

$$\begin{aligned}&\theta(t)=\hat{\theta}(t)+S(t),\\ &\dot{\hat{\theta}}(t)=KM(s_k)y(s_k)=KM(s_k)\Big[Q^*+\tfrac{1}{2}\left(\hat{\theta}(s_k)+S(s_k)-\theta^*\right)^T\\ &\quad\times H\left(\hat{\theta}(s_k)+S(s_k)-\theta^*\right)\Big],\quad t\in[t_k,t_{k+1}),\end{aligned} \tag{5}$$

where $\hat{\theta}_i(s)\in\left[\underline{\theta}_i^*,\overline{\theta}_i^*\right], s\le 0, i=1,2,$

$$\begin{aligned}&K=\begin{bmatrix}k_1 & 0\\ 0 & k_2\end{bmatrix},\quad S(t)=\left[a_1\mathrm{sq}\left(\tfrac{tl_1}{\varepsilon}\right),a_2\mathrm{sq}\left(\tfrac{tl_2}{\varepsilon}\right)\right]^T,\\ &M(t)=\left[\tfrac{1}{a_1}\mathrm{sq}\left(\tfrac{tl_1}{\varepsilon}\right),\tfrac{1}{a_2}\mathrm{sq}\left(\tfrac{tl_2}{\varepsilon}\right)\right]^T,\quad l_1=1,\quad l_2=2,\end{aligned} \tag{6}$$

in which $k_1>0$ and $k_2>0$ if $H<0$, or $k_1<0$ and $k_2<0$ if $H>0$, $\{s_k\}$ and $\{t_k\}$ denote the sampling and control update instants, respectively, which satisfy

$$\begin{aligned}&0=s_0<s_1<\ldots<s_k<\ldots,\quad \lim_{k\to\infty}s_k=\infty,\quad k\in\mathbb{Z}_0^+,\\ &s_{k+1}-s_k=\tfrac{1}{2}\tfrac{\varepsilon}{l_2}=\tfrac{1}{4}\tfrac{\varepsilon}{l_1}=\tfrac{\varepsilon}{4},\quad t_k=s_k+\eta\end{aligned} \tag{7}$$

with η being a constant delay. Note that in the case of continuous-time ES with sine wave in [20], $M(t)=\left[\frac{2}{a_1}\sin\left(\frac{2\pi l_1}{\varepsilon}t\right),\frac{2}{a_2}\sin\left(\frac{2\pi l_2}{\varepsilon}t\right)\right]^T$ and the convergence gain of the time-delay system (44) in [20] is KH. In (6), if $M(t)=\left[\frac{2}{a_1}\mathrm{sq}\left(\frac{tl_1}{\varepsilon}\right),\frac{2}{a_2}\mathrm{sq}\left(\frac{tl_2}{\varepsilon}\right)\right]^T$, the convergence gain of (22) will be $2KH$. For notational simplicity, $M(t)$ is chosen as (6) to achieve the same convergence gain as that in [20].

Thus, the estimation error is governed by

$$\begin{aligned}&\dot{\tilde{\theta}}(t)=KM(s_k)\left[Q^*+\tfrac{1}{2}\left(\tilde{\theta}(s_k)+S(s_k)\right)^T H\left(\tilde{\theta}(s_k)+S(s_k)\right)\right]\\ &=KM(s_k)\Big[Q^*+\tfrac{1}{2}\tilde{\theta}^T(s_k)H\tilde{\theta}(s_k)+S^T(s_k)H\tilde{\theta}(s_k)\\ &\quad+\tfrac{1}{2}S^T(s_k)HS(s_k)\Big],\quad t\in[t_k,t_{k+1}).\end{aligned} \tag{8}$$

Following the time-delay approach to sampled-data control (see [3, Chapter 7]), denote

$$h(t)=t-s_k,\quad t\in[t_k,t_{k+1}),\quad \eta\le h(t)<\eta+\tfrac{\varepsilon}{4}=h_M. \tag{9}$$

Taking into account

$$\begin{aligned}&\mathrm{sq}\left(\tfrac{s_kl_i}{\varepsilon}\right)=\mathrm{sq}\left(\tfrac{(t-\eta)l_i}{\varepsilon}\right),\quad t\in[t_k,t_{k+1}),\\ &\Longrightarrow\quad M(s_k)=M(t-\eta),\quad S(s_k)=S(t-\eta),\end{aligned} \tag{10}$$

the dynamics (8) becomes

$$\begin{aligned}\dot{\tilde{\theta}}(t)=KM(t-\eta)\Big[&S^T(t-\eta)H\tilde{\theta}(t-h(t))\\ &+\tfrac{1}{2}\tilde{\theta}^T(t-h(t))H\tilde{\theta}(t-h(t))\Big]+\omega(t),\end{aligned} \tag{11}$$

where

$$\omega(t)=KM(t-\eta)\left[Q^*+\tfrac{1}{2}S^T(t-\eta)HS(t-\eta)\right] \tag{12}$$

which satisfies

$$\begin{aligned}&|\omega(t)|\le|KM(t-\eta)|\left[|Q^*|+\tfrac{\lambda(H)}{2}S^2(t-\eta)\right]\\ &=\sqrt{\left(\tfrac{k_1^2}{a_1^2}+\tfrac{k_2^2}{a_2^2}\right)}\left(|Q^*|+\tfrac{\lambda(H)}{2}\left(a_1^2+a_2^2\right)\right)=\bar{\omega},\end{aligned} \tag{13}$$

with $\lambda(H)=\max\{|\lambda_{\max}(H)|,|\lambda_{\min}(H)|\}$.

B. A Time-delay Approach to Averaging

Based on [3, Chapter 7], the dynamics of ES under sampling (8) has been modeled as a typical time-delay plant (11) in which the time-varying delay $h(t)$ of (9) consists of the transmission delay η and sampling-induced delay bounded by $\frac{\varepsilon}{4}$. When there is no sampling and no delay, as well as the dither is of the form of sine wave, the conventional method for the stability analysis of (11) usually resorts to the averaged system via the averaging theorem [8]. To be specific, setting $h(t)=\eta\equiv 0$ and treating $\tilde{\theta}(t)$ as a "freeze" constant in the averaging analysis, we derive the averaged system of (11) as follows (see details in [20]):

$$\dot{\tilde{\theta}}_{av}(t)=KH\tilde{\theta}_{av}(t), \tag{14}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3140259, IEEE Transactions on Automatic Control

which is exponentially stable since KH is Hurwitz. Without losing of generality, assuming $H>0$, the controller gain is chosen as $K<0$. Thus, $\det(\lambda I-KH)=\begin{vmatrix}\lambda-k_1h_{11} & -k_1h_{12}\\ -k_2h_{12} & \lambda-k_2h_{22}\end{vmatrix}=\lambda^2-(k_1h_{11}+k_2h_{22})\lambda+k_1k_2\left(h_{11}h_{22}-h_{12}^2\right)=0$. The eigenvalues satisfy $\begin{cases}\lambda_1+\lambda_2=(k_1h_{11}+k_2h_{22})<0\\ \lambda_1\lambda_2=k_1k_2\left(h_{11}h_{22}-h_{12}^2\right)>0\end{cases}\Rightarrow\lambda_1<0,\lambda_2<0$, which proves KH is Hurwitz.

As clarified in [8, Chapter 10.4], the essential problem in the averaging method is to determine in what sense the behavior of the averaged system (14) approximates the behavior of the original system (11), which may not be intuitively clear. Averaging results from [8] are restricted to systems described by differential equations with continuous right-hand sides, that are not applicable to discontinuous square wave dithers. There are averaging results that are suitable for differential equations with a discontinuous right-hand side (as in [13], [16]), but they did not consider delays. What is more important, when the sampling and the delays are taken into account, the time-delay plant (11) is potentially unstable if the delay is large. Considering (9), the existing methods for ES have not given upper bounds on ε and η to ensure the stability of (11). Now the designer has a chance to know how large the sampling interval ε and the delay η that the ES control system is able to withstand through our time-delay approach.

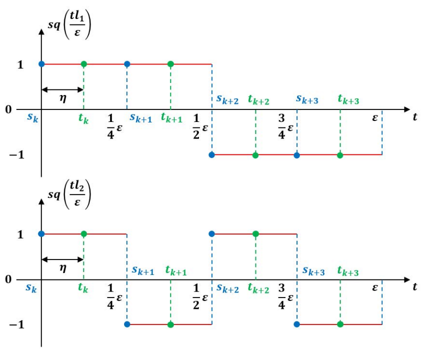


Fig. 3. The distinct square waves of two dithers.

Note that, the equation (11) is a system with time-varying parameters and delays whose stability is hard to judge at this stage. To overcome this difficulty, each element of (11) is expanded as

$$\begin{aligned}\dot{\tilde{\theta}}_i(t)&=k_i\left(h_{ii}\tilde{\theta}_i(t-h(t))+h_{ij}\tilde{\theta}_j(t-h(t))\right)\\&+\frac{k_ia_j}{a_i}\mathrm{sq}\left(\frac{(t-\eta)l_i}{\varepsilon}\right)\mathrm{sq}\left(\frac{(t-\eta)l_j}{\varepsilon}\right)\\&\times\left(h_{ij}\tilde{\theta}_i(t-h(t))+h_{jj}\tilde{\theta}_j(t-h(t))\right)+\frac{k_i}{a_i}\mathrm{sq}\left(\frac{(t-\eta)l_i}{\varepsilon}\right)Q^*\\&+\frac{k_i}{2a_i}\mathrm{sq}\left(\frac{(t-\eta)l_i}{\varepsilon}\right)\tilde{\theta}^T(t-h(t))H\tilde{\theta}(t-h(t))\\&+\frac{k_i}{2a_i}\mathrm{sq}\left(\frac{(t-\eta)l_i}{\varepsilon}\right)S^T(t-\eta)HS(t-\eta),\quad i,j=1,2,i\neq j.\end{aligned}\tag{15}$$

As illustrated in Fig. 3, the sampling is periodic with the sampling interval in (7) equal to one half of the period of the dither $\mathrm{sq}\left(\frac{tl_2}{\varepsilon}\right)$ and one fourth of the period of the dither $\mathrm{sq}\left(\frac{tl_1}{\varepsilon}\right)$. Thus we obtain

$$\begin{aligned}&\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)d\tau=0,\\&\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\mathrm{sq}\left(\frac{(\tau-\eta)l_j}{\varepsilon}\right)d\tau=0,\\&\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)S^T(\tau-\eta)HS(\tau-\eta)d\tau\\&=\int_{t-\varepsilon}^t\Big[h_{ii}a_i^2\mathrm{sq}^3\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)+2h_{ij}a_1a_2\mathrm{sq}^2\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\\&\quad\times\mathrm{sq}\left(\frac{(\tau-\eta)l_j}{\varepsilon}\right)+h_{jj}a_j^2\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\mathrm{sq}^2\left(\frac{(\tau-\eta)l_j}{\varepsilon}\right)\Big]d\tau=0.\end{aligned}\tag{16}$$

Following [4], we apply the time-delay method to averaging of (15). Integrating (15) backward in $t\geq\varepsilon+\eta$ from $t-\varepsilon$ to t (by "backward", we refer to the integral interval $[t-\varepsilon,t]$ rather than $[t,t+\varepsilon]$ which was used in [8]), and taking into account (16), we get

$$\begin{aligned}&\frac{1}{\varepsilon}\int_{t-\varepsilon}^t\dot{\tilde{\theta}}_i(\tau)d\tau=\frac{k_i}{\varepsilon}\int_{t-\varepsilon}^t\left(h_{ii}\tilde{\theta}_i(\tau-h(\tau))+h_{ij}\tilde{\theta}_j(\tau-h(\tau))\right)d\tau\\&\quad+\frac{k_ia_j}{a_i\varepsilon}\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\mathrm{sq}\left(\frac{(\tau-\eta)l_j}{\varepsilon}\right)\\&\quad\times\left(h_{ij}\tilde{\theta}_i(\tau-h(\tau))+h_{jj}\tilde{\theta}_j(\tau-h(\tau))\right)d\tau\\&\quad+\frac{k_i}{2a_i\varepsilon}\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\tilde{\theta}^T(\tau-h(\tau))H\tilde{\theta}(\tau-h(\tau))d\tau.\end{aligned}\tag{17}$$

To handle the 1st term on the right-hand side of (17), we have

$$\begin{aligned}&\frac{k_i}{\varepsilon}\int_{t-\varepsilon}^t\left(h_{ii}\tilde{\theta}_i(\tau-h(\tau))+h_{ij}\tilde{\theta}_j(\tau-h(\tau))\right)d\tau\\&=\frac{k_i}{\varepsilon}\int_{t-\varepsilon}^t\Big[\left(h_{ii}\tilde{\theta}_i(\tau-h(\tau))+h_{ij}\tilde{\theta}_j(\tau-h(\tau))\right)\\&\quad\pm\left(h_{ii}\tilde{\theta}_i(t)+h_{ij}\tilde{\theta}_j(t)\right)\Big]d\tau\\&=k_i\left(h_{ii}\tilde{\theta}_i(t)+h_{ij}\tilde{\theta}_j(t)\right)-\frac{k_i}{\varepsilon}\int_{t-\varepsilon}^t\Big[\left(h_{ii}\tilde{\theta}_i(t)+h_{ij}\tilde{\theta}_j(t)\right)\\&\quad-\left(h_{ii}\tilde{\theta}_i(\tau-h(\tau))+h_{ij}\tilde{\theta}_j(\tau-h(\tau))\right)\Big]d\tau\\&=k_i\left(h_{ii}\tilde{\theta}_i(t)+h_{ij}\tilde{\theta}_j(t)\right)\\&\quad-\frac{k_i}{\varepsilon}\int_{t-\varepsilon}^t\int_{\tau-h(\tau)}^t\left(h_{ii}\dot{\tilde{\theta}}_i(s)+h_{ij}\dot{\tilde{\theta}}_j(s)\right)dsd\tau.\end{aligned}\tag{18}$$

To address the 2nd term on the right-hand side of (17), we get

$$\begin{aligned}&\frac{k_ia_j}{a_i\varepsilon}\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\mathrm{sq}\left(\frac{(\tau-\eta)l_j}{\varepsilon}\right)\\&\quad\times\left(h_{ij}\tilde{\theta}_i(\tau-h(\tau))+h_{jj}\tilde{\theta}_j(\tau-h(\tau))\right)d\tau\\&=-\frac{k_ia_j}{a_i\varepsilon}\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\mathrm{sq}\left(\frac{(\tau-\eta)l_j}{\varepsilon}\right)\Big[\left(h_{ij}\tilde{\theta}_i(t)+h_{jj}\tilde{\theta}_j(t)\right)\\&\quad-\left(h_{ij}\tilde{\theta}_i(\tau-h(\tau))+h_{jj}\tilde{\theta}_j(\tau-h(\tau))\right)\Big]d\tau\\&=-\frac{k_ia_j}{a_i\varepsilon}\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\mathrm{sq}\left(\frac{(\tau-\eta)l_j}{\varepsilon}\right)\\&\quad\times\int_{\tau-h(\tau)}^t\left(h_{ij}\dot{\tilde{\theta}}_i(s)+h_{jj}\dot{\tilde{\theta}}_j(s)\right)dsd\tau,\end{aligned}\tag{19}$$

where we utilize (16). To handle the 3rd term on the right-hand side of (17), we have

$$\begin{aligned}&\frac{k_i}{2a_i\varepsilon}\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\tilde{\theta}^T(\tau-h(\tau))H\tilde{\theta}(\tau-h(\tau))d\tau\\&=-\frac{k_i}{2a_i\varepsilon}\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\Big[\tilde{\theta}^T(t)H\tilde{\theta}(t)\\&\quad-\tilde{\theta}^T(\tau-h(\tau))H\tilde{\theta}(\tau-h(\tau))\Big]d\tau\\&=-\frac{k_i}{a_i\varepsilon}\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\int_{\tau-h(\tau)}^t\tilde{\theta}^T(s)H\dot{\tilde{\theta}}(s)dsd\tau.\end{aligned}\tag{20}$$

Substituting (18), (19) and (20) into (17), we obtain

$$\begin{aligned}&\frac{1}{\varepsilon}\int_{t-\varepsilon}^t\dot{\tilde{\theta}}_i(\tau)d\tau=k_i\left(h_{ii}\tilde{\theta}_i(t)+h_{ij}\tilde{\theta}_j(t)\right)\\&\quad-\frac{k_i}{\varepsilon}\int_{t-\varepsilon}^t\int_{\tau-h(\tau)}^t\left(h_{ii}\dot{\tilde{\theta}}_i(s)+h_{ij}\dot{\tilde{\theta}}_j(s)\right)dsd\tau\\&\quad-\frac{k_ia_j}{a_i\varepsilon}\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\mathrm{sq}\left(\frac{(\tau-\eta)l_j}{\varepsilon}\right)\\&\quad\times\int_{\tau-h(\tau)}^t\left(h_{ij}\dot{\tilde{\theta}}_i(s)+h_{jj}\dot{\tilde{\theta}}_j(s)\right)dsd\tau\\&\quad-\frac{k_i}{a_i\varepsilon}\int_{t-\varepsilon}^t\mathrm{sq}\left(\frac{(\tau-\eta)l_i}{\varepsilon}\right)\int_{\tau-h(\tau)}^t\tilde{\theta}^T(s)H\dot{\tilde{\theta}}(s)dsd\tau.\end{aligned}\tag{21}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3140259, IEEE Transactions on Automatic Control

Rearranging (21) into a vector form, we arrive at the closed-loop system

$$\frac{d}{dt}\left[\tilde{\theta}(t)-G(t)\right]=KH\tilde{\theta}(t)-KHY_1(t)-\bar{K}HY_0(t)-\bar{\bar{K}}Y_2(t),\quad t\geq\varepsilon+\eta, \tag{22}$$

where

$$\begin{aligned}
&G(t)=\frac{1}{\varepsilon}\int_{t-\varepsilon}^{t}(\tau-t+\varepsilon)\dot{\tilde{\theta}}(\tau)d\tau,\\
&Y_1(t)=\frac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau-h(\tau)}^{t}\dot{\tilde{\theta}}(s)dsd\tau\\
&=\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^{t}\dot{\tilde{\theta}}(s)dsd\zeta,\\
&Y_0(t)=\frac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau-h(\tau)}^{t}N_1(\tau)\dot{\tilde{\theta}}(s)dsd\tau\\
&=\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^{t}N_1(t-\varepsilon\zeta)\dot{\tilde{\theta}}(s)dsd\zeta,\\
&Y_2(t)=\frac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau-h(\tau)}^{t}N_2(\tau)\tilde{\theta}^T(s)H\dot{\tilde{\theta}}(s)dsd\tau\\
&=\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^{t}N_2(t-\varepsilon\zeta)\tilde{\theta}^T(s)H\dot{\tilde{\theta}}(s)dsd\zeta,
\end{aligned} \tag{23}$$

with the change of variable $\varepsilon\zeta=t-\tau$ and

$$\bar{K}=\begin{bmatrix}0 & \frac{k_1a_2}{a_1}\\ \frac{k_2a_1}{a_2} & 0\end{bmatrix},\quad \bar{\bar{K}}=\begin{bmatrix}\frac{k_1}{a_1} & 0\\ 0 & \frac{k_2}{a_2}\end{bmatrix},\quad N_2(\tau)=\begin{bmatrix}\mathrm{sq}\left(\frac{(\tau-\eta)l_1}{\varepsilon}\right)\\ \mathrm{sq}\left(\frac{(\tau-\eta)l_2}{\varepsilon}\right)\end{bmatrix}$$
$$N_1(\tau)=\mathrm{sq}\left(\frac{(\tau-\eta)l_1}{\varepsilon}\right)\mathrm{sq}\left(\frac{(\tau-\eta)l_2}{\varepsilon}\right). \tag{24}$$

C. Closed-loop Solutions without Approximations

Note that $\dot{\tilde{\theta}}(t)$ in (22)-(23) is given by (11). If we substitute the right-hand side of (11) into (23), we arrive at a differential equation with distributed delays. If $\tilde{\theta}(t)$ and $\dot{\tilde{\theta}}(t)$ are of order $O(1)$, and the delay η is of order $O(\varepsilon)$, then the integral terms $G(t),Y_1(t),Y_2(t)$ are of order $O(\varepsilon)$ and are close to zero when ε is chosen to be sufficiently small. At this stage, the ES system (11) has been further transformed into the time-delay system (22) with $\dot{\tilde{\theta}}(t)$ defined by (11) when $t\geq\varepsilon+\eta$.

Looking at the averaged system (14) of the original ES system (11) and the transformed time-delay system (22), it is apparent that the averaged system and the time-delay system have the consistent dominant part $\frac{d}{dt}\tilde{\theta}(t)=KH\tilde{\theta}(t)$ which is stable (KH is Hurwitz) and delay-free. The differences are the additional terms with distributed delays $G(t),Y_1(t),Y_2(t)$ that vanish for $\varepsilon\to0,\eta\to0$. That is to say, to describe the behavior of the original ES system (11), the time-delay plant (22) is an accurate model as an alternative to the stable averaged system (14), with explicit perturbation terms $G(t),Y_1(t),Y_2(t)$ in (23). When the sampling interval ε and the delay η are large, the perturbed impact of $G(t),Y_1(t),Y_2(t)$ on the stability of (22) is strong. Thus, to find the upper bounds on ε and η to preserve the practical stability of (22) is important. For the definition of practical stability see e.g. [16, Section 2.2].

An accurate representation with quantitative bounds on ε, η and the resulting ultimate bound of the closed-loop ES solution is the major advantage of the developed time-delay method. The traditional classical averaging, weak limit averaging and Lie bracket methods are "approximate" methods in nature:

- the classical averaging method presents the solution of the original ES system as an approximation by the solution of the averaged system [1], [9].
- the Lie bracket method employs the averaged system written in terms of Lie brackets to approximate the behavior of the original ES control system [2], [16].

Different from the two above methods, the proposed time-delay approach is a direct method without any approximation. The solution $\tilde{\theta}(t)$ of the ES system (11) is also a solution of the time-delay system (22). Thus, the stability of the time-delay system guarantees the stability of the original ES control system.

D. LMI-based Ultimate Boundedness of the Error System

To formulate the main theorem, we define the following LMIs that depend on the tuning parameters k_1,k_2,a_1,a_2 and $q,\delta,\varepsilon^*,\eta^*>0$ $\left(\text{with } h_M^*=\eta^*+\frac{\varepsilon^*}{4}\right)$ as well as $\sigma>\sigma_0$, and the following decision scalars $q_2,b,\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_P,\lambda_R,\lambda_{Q_0},\lambda_{Q_1},\lambda_{S_0},\lambda_{R_0},\lambda_W>0$, and 2×2 matrices $P>I$ and $R,Q_0,Q_1,S_0,R_0,W>0$. Consider the LMIs:

$$\begin{aligned}
&\Phi_1=\begin{bmatrix}P-I & -P\\ * & P+e^{-2\delta\varepsilon^*}R\end{bmatrix}>0,\quad \Phi_2=\begin{bmatrix}\Omega & \Psi_\varepsilon & \Psi_\eta\\ * & -\frac{1}{\varepsilon^*}\Xi_\varepsilon & 0\\ * & * & -\frac{1}{\eta^*}\Xi_\eta\end{bmatrix}<0,\\
&\Phi_3=(\sigma_0+(\varepsilon^*+\eta^*)\bar{\omega})^2e^{2\mu(\varepsilon^*+\eta^*)}<\sigma^2,\\
&\Phi_4=\left(1+\frac{1}{q}\right)\lambda_P(\sigma_0+(\varepsilon^*+\eta^*)\bar{\omega})^2e^{2\mu(\varepsilon^*+\eta^*)}+\eta^*\lambda_{S_0}\sigma^2\\
&\quad+\Big[\frac{\varepsilon^{*2}(1+q)}{4}\lambda_P+\frac{\varepsilon^{*2}}{3}\lambda_R+\frac{\left(\varepsilon^{*2}+3\varepsilon^*h_M^*+3h_M^{*2}\right)}{3}\\
&\quad\times\left(\lambda_{Q_1}+\lambda_{Q_0}+q_2\lambda_{\max}(H^2)\sigma^2\right)+\frac{1}{2}\eta^{*3}\lambda_{R_0}\\
&\quad+\frac{h_M^*\varepsilon^{*2}}{16}e^{\frac{\delta\varepsilon^*}{2}}\lambda_W\Big]\Delta^2+\frac{\varepsilon^*+\eta^*}{2\delta}b\bar{\omega}^2<\sigma^2,\\
&P-\lambda_PI<0,\quad R-\lambda_RI<0,\quad Q_0-\lambda_{Q_0}I<0,\\
&Q_1-\lambda_{Q_1}I<0,\quad S_0-\lambda_{S_0}I<0,\quad R_0-\lambda_{R_0}I<0,\\
&W-\lambda_WI<0,
\end{aligned} \tag{25}$$

where Ω is composed of

$$\begin{aligned}
&\Omega_{11}=(PKH+H^TK^TP+2\delta PI)+S_0-e^{-2\delta\eta^*}R_0,\\
&\Omega_{12}=-(H^TK^T+2\delta I)P,\quad \Omega_{13}=-PKH,\\
&\Omega_{14}=-P\bar{K}H,\quad \Omega_{15}=-P\bar{\bar{K}},\quad \Omega_{16}=e^{-2\delta\eta^*}R_0,\\
&\Omega_{22}=-\frac{4}{\varepsilon^*}e^{-2\delta\varepsilon^*}R+2\delta P,\quad \Omega_{23}=PKH,\quad \Omega_{24}=P\bar{K}H,\\
&\Omega_{25}=P\bar{\bar{K}},\quad \Omega_{33}=-\frac{4}{\varepsilon^*+2h_M^*}e^{-2\delta(\varepsilon^*+h_M^*)}Q_1,\\
&\Omega_{44}=-\frac{4}{\varepsilon^*+2h_M^*}e^{-2\delta(\varepsilon^*+h_M^*)}Q_0,\\
&\Omega_{55}=-\frac{2q_2}{\varepsilon^*+2h_M^*}e^{-2\delta(\varepsilon^*+h_M^*)}I,\\
&\Omega_{66}=-e^{-2\delta\eta^*}(S_0+R_0)+\left[\lambda_1+2\lambda_3\lambda^2(H)\sigma^2\right]I,\\
&\Omega_{77}=-\frac{\pi^2}{4}e^{-2\delta\eta^*}W+\left[\lambda_2+8\lambda_4\lambda_{\max}(H^2)\sigma^2\right]I,\\
&\Omega_{88}=-\lambda_1I,\quad \Omega_{99}=-\lambda_2I,\quad \Omega_{10,10}=-\lambda_3I,\\
&\Omega_{11,11}=-\lambda_4I,\quad \Omega_{12,12}=-(\varepsilon^*+\eta^*)bI,
\end{aligned} \tag{26}$$

and $\bar{\omega}$ is given by (13),

$$\begin{aligned}
&\mu=\sqrt{\left(\frac{k_1^2}{a_1^2}+\frac{k_2^2}{a_2^2}\right)}\left(\sqrt{\lambda_{\max}(H^2)\left(a_1^2+a_2^2\right)}+\frac{\lambda(H)}{2}\sigma\right),\\
&\Xi_\varepsilon=R+\frac{3}{2}\left(Q_1+Q_0+q_2\sigma^2H^2\right)+\frac{\varepsilon^*}{16}e^{\frac{\delta\varepsilon^*}{2}}W,\\
&\Xi_\eta=2\left(Q_1+Q_0+q_2\sigma^2H^2\right)+\eta^*R_0,
\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3140259, IEEE Transactions on Automatic Control

$$\Psi_\varepsilon = \left[0,0,0,0,0,KH,KH,\bar{K}H,\bar{K}H,\frac{\bar{\bar{K}}}{2},\frac{\bar{\bar{K}}}{2},I\right]^T \Xi_\varepsilon,$$
$$\Psi_\eta = \left[0,0,0,0,0,KH,KH,\bar{K}H,\bar{K}H,\frac{\bar{\bar{K}}}{2},\frac{\bar{K}}{2},I\right]^T \Xi_\eta,$$
$$\Delta = \sqrt{\frac{k_1^2}{a_1^2}+\frac{k_2^2}{a_2^2}}\Big[|Q^*| + \frac{\lambda(H)}{2}\Big((\sigma_0+(\varepsilon+\eta)\bar{\omega})e^{\mu(\varepsilon+\eta)} + \sqrt{a_1^2+a_2^2}\Big)^2\Big]. \quad (27)$$

Theorem 1: Assume that the Hessian H and the extremum value Q^* are known, the extremum point θ^* is uncertain but each of its elements belongs to a known ball $\theta_i^* \in \left[\underline{\theta}_i^*, \overline{\theta}_i^*\right], i=1,2$. Consider the closed-loop system consisting of the multi-input plant (1) and the ES controller (5), with the initial condition satisfying $\left|\tilde{\theta}(0)\right| \le \sigma_0$. Given the tuning parameters $k_1,k_2,a_1,a_2,q,\delta,\varepsilon^*,\eta^*,\sigma,\sigma_0$, let the scalars $q_2,b,\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_P,\lambda_R,\lambda_{Q_0},\lambda_{Q_1},\lambda_{S_0},\lambda_{R_0},\lambda_W$, and the matrices P,R,Q_0,Q_1,S_0,R_0,W satisfy the LMIs (25) with notations given by (26), (27). Then, $\forall\varepsilon \in (0,\varepsilon^*]$ and $\forall\eta \in (0,\eta^*]$, the solution of the closed-loop system (11) satisfies

$$\begin{aligned}
&\left|\tilde{\theta}(t)\right|^2 < \left(\left|\tilde{\theta}(0)\right| + (\varepsilon+\eta)\bar{\omega}\right)^2 e^{2\mu t} < \sigma^2, \quad t\in[0,\varepsilon+\eta],\\
&\left|\tilde{\theta}(t)\right|^2 < \left(1+\tfrac{1}{q}\right)\lambda_P e^{-2\delta(t-\varepsilon-\eta)}\left(\left|\tilde{\theta}(0)\right| + (\varepsilon+\eta)\bar{\omega}\right)^2\\
&\times e^{2\mu(\varepsilon+\eta)} + e^{-2\delta(t-\varepsilon-\eta)}\eta\lambda_{S_0}\sigma^2\\
&+e^{-2\delta(t-\varepsilon-\eta)}\Big[\tfrac{\varepsilon^2(1+q)}{4}\lambda_P + \tfrac{\varepsilon^2}{3}\lambda_R\\
&+\tfrac{(\varepsilon^2+3\varepsilon h_M+3h_M^2)}{3}\left(\lambda_{Q_1}+\lambda_{Q_0}+q_2\lambda_{\max}(H^2)\sigma^2\right)\\
&+\tfrac{1}{2}\eta^3\lambda_{R_0} + \tfrac{h_M\varepsilon^2}{16}e^{\frac{\delta\varepsilon}{2}}\lambda_W\Big]\Delta^2\\
&+\left(1-e^{-2\delta(t-\varepsilon-\eta)}\right)\tfrac{\varepsilon+\eta}{2\delta}b\bar{\omega}^2 < \sigma^2, \quad t\ge\varepsilon+\eta,
\end{aligned} \quad (28)$$

Moreover, the solution exponentially approaches the ball

$$\Theta = \left\{\tilde{\theta}\in\mathbb{R}^2 : \left|\tilde{\theta}\right|^2 < \tfrac{\varepsilon+\eta}{2\delta}b\bar{\omega}^2\right\}. \quad (29)$$

with a decay rate δ.

Proof: The proof is given in Appendix. ■

Comparing the LKF (49) with the LKF (A.6) in [20] that included only $V_P(t)$ and $V_R(t)$ terms, it is observed that the LKF in (49) also contains $V_{Q_1}(t)$ and $V_{Q_2}(t)$ which are employed to compensate the linear perturbation $Y_1(t)$ and the nonlinear perturbation $Y_2(t)$. In [20], $Y_1(t)$ and $Y_2(t)$ are treated as disturbances. As a result, the ultimate bound achieved with the improved LKF of this paper is smaller than that in [20] (see the improvement in the example of [21]).

E. Selection of ES Parameters and Uncertainties

The ultimate bound on the error of extremum seeking depends upon both the sampling interval ε and the delay η from (29). It means the smaller ε and η suggest the less deviation of the estimate of the extremum point from its real value. Applying Schur complement to $\Phi_1 > 0$ first, we have $P - I - P\left(P+e^{-2\delta\varepsilon^*}R\right)^{-1}P^T \approx P - I - P(P+R)^{-1}P^T > 0$, when $\varepsilon^* \to 0$. Since P and R are decision variables that are mutually independent, the above inequality is always feasible. Given P and R subject to $\Phi_1 > 0$, applying Schur complement to $\Phi_2 < 0$, we get $\Omega + \varepsilon^*\Psi_\varepsilon\Xi_\varepsilon{}^{-1}\Psi_\varepsilon^T + \eta^*\Psi_\eta\Xi_\eta{}^{-1}\Psi_\eta^T \approx \Omega < 0$, when $\varepsilon^*\to 0, \eta^*\to 0$. Further applying Schur complement to $\Omega < 0$, there is no difficulty to prove that $\Phi_2 < 0$ always hold for $\varepsilon^*\to 0, \eta^*\to 0$. The feasibility of $\Phi_3 < \sigma^2$ and $\Phi_4 < \sigma^2$ are self-evident. Above all, provided that ε and η are sufficiently small, the LMIs (25) are always feasible. That is to say, given any pair of the initial condition and overall bound satisfying $|\tilde{\theta}(0)| \le \sigma_0 < \sigma$, as long as the dither and sampling period ε and the delay η do not go beyond the interval $(0,\varepsilon^*]$ and $(0,\eta^*]$, the ultimate bound (29) is available and could be reduced by decreasing ε if η is of order $O(\varepsilon)$. Thus, the practical stability is semi-global.

TABLE I
TWO-VARIABLE SYSTEM ($Q^* = 0, H = \mathrm{diag}\{2,2\}$)

	ε^*	η^*	σ_0	σ	δ	UB
Continuous-time ES: sine wave						
[20]	0.14	-	0.1	$\sqrt{2}$	0.01	1.32
[20]	0.017	-	$\sqrt{2}$	$2\sqrt{2}$	0.01	1.90
Sampled-data ES: square wave						
Theorem 1	0.24	0.1	0.1	$\sqrt{2}$	0.01	0.65
Theorem 1	0.09	0.01	$\sqrt{2}$	$2\sqrt{2}$	0.01	0.18

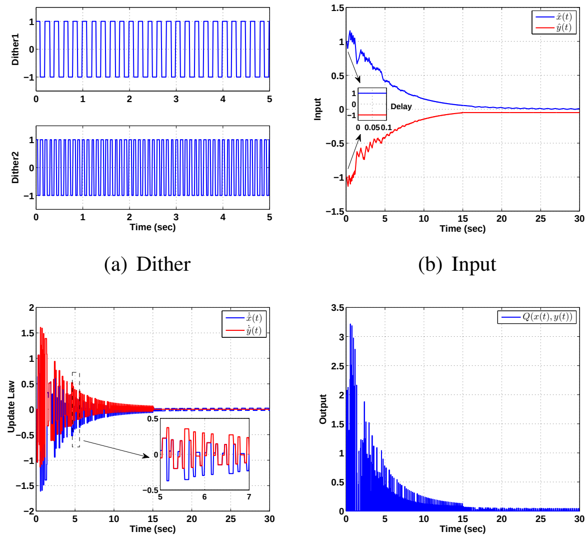
In classical ES, the Hessian H, the extremum value Q^* and the extremum point θ^* in (1) are assumed to be unknown. In the face of an unknown "black box" model, it is impossible to choose tuning parameters and even to perform simulations. Here we study a "grey box", where H is known, Q^* is known up to small uncertainties (see Corollary 2 of [21]), θ^* is uncertain but belongs to a known ball. A quantitative analysis is summarized in Theorem 1. When H is also uncertain (norm-bounded type or polytopic type), the derivation of LMI condition follows the logic of scalar case (see Corollary 2 of [21]) but it is more complex and we do not go into details due to page limit. There is a trade-off between the quantitative analysis with the plant information and the qualitative analysis without the model knowledge.

III. NUMERICAL EXAMPLE

In this section we consider an autonomous vehicle in an environment without GPS orientation [16]. The goal is to reach the location of the stationary minimum of the map $Q(x(t),y(t)) = Q^* + \frac{1}{2}[x(t),y(t)]H\begin{bmatrix}x(t)\\y(t)\end{bmatrix} = x^2(t)+y^2(t)$, where $Q^*=0$ and $H=\begin{bmatrix}2&0\\0&2\end{bmatrix}$ are known. Note that here the notation $[x(t),y(t)]^T$ (which is consistent with those in [16]) refers to a vehicle's trajectory in a plane with x-axis and y-axis. It is equivalent to the input vector $\theta(t) = [\theta_1(t),\theta_2(t)]^T$ in (1)-(2). The ES parameters are chosen as $k_1=k_2=-0.01, a_1=a_2=0.2$. The LMI solution is shown in Table I. The numerical simulations are shown in Fig. 4, where $\varepsilon=0.2$ and $\eta=0.1$ which are larger than those achieved via LMIs. This shows conservatism of results of LMIs.

IV. CONCLUSIONS

In this article we offer a constructive method based on the time-delay approach to averaging and to sampled-data control for the design and analysis of ES with discontinuous dither and constant delay. The resulting time-delay method allows to derive quantitative bounds on the extremum seeking error, the period of the dither and sampling, as well as on the largest



(a) Dither (b) Input

(c) Update Law (d) Output

Fig. 4. Numerical simulation of sampled-data ES.

delay that the ES system is able to tolerate, provided the extremum value and the extremum point are uncertain.

APPENDIX-PROOF OF THEOREM 1

The proof is divided into two steps.

- Step 1—we prove the practical stability of system (22). We employ

$$V_P(t) = \left[\tilde{\theta}(t) - G(t)\right]^T P\left[\tilde{\theta}(t) - G(t)\right]. \tag{30}$$

Then, we have

$$\begin{aligned}
&\dot{V}_P(t) + 2\delta V_P(t) = 2\left[\tilde{\theta}(t) - G(t)\right]^T P[KH\tilde{\theta}(t) - KHY_1(t) \\
&\quad -\bar{K}HY_0(t) - \bar{\bar{K}}Y_2(t)] + 2\delta\left[\tilde{\theta}(t) - G(t)\right]^T P\left[\tilde{\theta}(t) - G(t)\right] \\
&= 2\tilde{\theta}^T(t)P(KH + \delta I)\tilde{\theta}(t) + 2\delta G^T(t)PG(t) \\
&\quad -2\tilde{\theta}^T(t)(H^TK^T + 2\delta I)PG(t) \\
&\quad -2\tilde{\theta}^T(t)PKHY_1(t) - 2\tilde{\theta}^T(t)P\bar{K}HY_0(t) - 2\tilde{\theta}^T(t)P\bar{\bar{K}}Y_2(t) \\
&\quad +2G^T(t)PKHY_1(t) + 2G^T(t)P\bar{K}HY_0(t) + 2G^T(t)P\bar{\bar{K}}Y_2(t).
\end{aligned} \tag{31}$$

To compensate $G(t)$ in (31), we use the following:

$$V_R(t) = \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t} e^{-2\delta(t-\tau)}(\tau - t + \varepsilon)^2\dot{\tilde{\theta}}^T(\tau)R\dot{\tilde{\theta}}(\tau)d\tau. \tag{32}$$

Hence, we get

$$\begin{aligned}
&\dot{V}_R(t) + 2\delta V_R(t) = \varepsilon\dot{\tilde{\theta}}^T(t)R\dot{\tilde{\theta}}(t) \\
&\quad -\tfrac{2}{\varepsilon}\int_{t-\varepsilon}^{t} e^{-2\delta(t-\tau)}(\tau - t + \varepsilon)\dot{\tilde{\theta}}^T(\tau)R\dot{\tilde{\theta}}(\tau)d\tau \\
&\leq \varepsilon\dot{\tilde{\theta}}^T(t)R\dot{\tilde{\theta}}(t) - \tfrac{2}{\varepsilon}e^{-2\delta\varepsilon}\int_{t-\varepsilon}^{t}(\tau - t + \varepsilon)\dot{\tilde{\theta}}^T(\tau)R\dot{\tilde{\theta}}(\tau)d\tau \\
&\leq \varepsilon\dot{\tilde{\theta}}^T(t)R\dot{\tilde{\theta}}(t) - \tfrac{4}{\varepsilon}e^{-2\delta\varepsilon}G^T(t)RG(t),
\end{aligned} \tag{33}$$

where the extended Jensen's inequality is used

$$\begin{aligned}
&2G^T(t)RG(t) = \tfrac{2}{\varepsilon^2}\int_{t-\varepsilon}^{t}(\tau - t + \varepsilon)\dot{\tilde{\theta}}^T(\tau)d\tau R \\
&\quad \times\int_{t-\varepsilon}^{t}(\tau - t + \varepsilon)\dot{\tilde{\theta}}(\tau)d\tau \\
&\leq \tfrac{2}{\varepsilon^2}\int_{t-\varepsilon}^{t}(\tau - t + \varepsilon)d\tau\cdot\int_{t-\varepsilon}^{t}(\tau - t + \varepsilon)\dot{\tilde{\theta}}^T(\tau)R\dot{\tilde{\theta}}(\tau)d\tau \\
&= \int_{t-\varepsilon}^{t}(\tau - t + \varepsilon)\dot{\tilde{\theta}}^T(\tau)R\dot{\tilde{\theta}}(\tau)d\tau.
\end{aligned} \tag{34}$$

To compensate $Y_1(t)$ in (31), we define

$$\begin{aligned}
V_{Q_1}(t) &= 2\int_0^1\int_{t-\varepsilon\zeta-h_M}^{t} e^{-2\delta(t-s)}(s - t + \varepsilon\zeta + h_M) \\
&\quad \times\dot{\tilde{\theta}}^T(s)Q_1\dot{\tilde{\theta}}(s)dsd\zeta.
\end{aligned} \tag{35}$$

Thus, we have

$$\begin{aligned}
&\dot{V}_{Q_1}(t) + 2\delta V_{Q_1}(t) = (\varepsilon + 2h_M)\dot{\tilde{\theta}}^T(t)Q_1\dot{\tilde{\theta}}(t) \\
&\quad -2\int_0^1\int_{t-\varepsilon\zeta-h_M}^{t} e^{-2\delta(t-s)}\dot{\tilde{\theta}}^T(s)Q_1\dot{\tilde{\theta}}(s)dsd\zeta \\
&\leq (\varepsilon + 2h_M)\dot{\tilde{\theta}}^T(t)Q_1\dot{\tilde{\theta}}(t) \\
&\quad -2e^{-2\delta(\varepsilon+h_M)}\int_0^1\int_{t-\varepsilon\zeta-h_M}^{t}\dot{\tilde{\theta}}^T(s)Q_1\dot{\tilde{\theta}}(s)dsd\zeta \\
&< (\varepsilon + 2h_M)\dot{\tilde{\theta}}^T(t)Q_1\dot{\tilde{\theta}}(t) - \tfrac{4}{\varepsilon+2h_M}e^{-2\delta(\varepsilon+h_M)}Y_1^T(t)Q_1Y_1(t),
\end{aligned} \tag{36}$$

where the extended Jensen's inequality is employed

$$\begin{aligned}
&2Y_1^T(t)Q_1Y_1(t) = 2\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^{t}\dot{\tilde{\theta}}^T(s)dsd\zeta \\
&\quad \times Q_1\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^{t}\dot{\tilde{\theta}}(s)dsd\zeta \\
&\leq 2\int_0^1(\varepsilon\zeta + h(t - \varepsilon\zeta))d\zeta \\
&\quad \times\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^{t}\dot{\tilde{\theta}}^T(s)Q_1\dot{\tilde{\theta}}(s)dsd\zeta \\
&< 2\int_0^1(\varepsilon\zeta + h_M)d\zeta\cdot\int_0^1\int_{t-\varepsilon\zeta-h_M}^{t}\dot{\tilde{\theta}}^T(s)Q_1\dot{\tilde{\theta}}(s)dsd\zeta \\
&= (\varepsilon + 2h_M)\int_0^1\int_{t-\varepsilon\zeta-h_M}^{t}\dot{\tilde{\theta}}^T(s)Q_1\dot{\tilde{\theta}}(s)dsd\zeta.
\end{aligned} \tag{37}$$

To compensate $Y_0(t)$ in (31), we define

$$\begin{aligned}
V_{Q_0}(t) &= 2\int_0^1\int_{t-\varepsilon\zeta-h_M}^{t} e^{-2\delta(t-s)}(s - t + \varepsilon\zeta + h_M) \\
&\quad \times\dot{\tilde{\theta}}^T(s)N_1^T(t - \varepsilon\zeta)Q_0N_1(t - \varepsilon\zeta)\dot{\tilde{\theta}}(s)dsd\zeta.
\end{aligned} \tag{38}$$

Consequently, we have

$$\begin{aligned}
&\dot{V}_{Q_0}(t) + 2\delta V_{Q_0}(t) = (\varepsilon + 2h_M)\dot{\tilde{\theta}}^T(t)Q_0\dot{\tilde{\theta}}(t) - 2\int_0^1\int_{t-\varepsilon\zeta-h_M}^{t} \\
&\quad \times e^{-2\delta(t-s)}\dot{\tilde{\theta}}^T(s)N_1^T(t - \varepsilon\zeta)Q_0N_1(t - \varepsilon\zeta)\dot{\tilde{\theta}}(s)dsd\zeta \\
&\leq (\varepsilon + 2h_M)\dot{\tilde{\theta}}^T(t)Q_0\dot{\tilde{\theta}}(t) - 2e^{-2\delta(\varepsilon+h_M)}\int_0^1\int_{t-\varepsilon\zeta-h_M}^{t} \\
&\quad \times\dot{\tilde{\theta}}^T(s)N_1^T(t - \varepsilon\zeta)Q_0N_1(t - \varepsilon\zeta)\dot{\tilde{\theta}}(s)dsd\zeta \\
&< (\varepsilon + 2h_M)\dot{\tilde{\theta}}^T(t)Q_0\dot{\tilde{\theta}}(t) - \tfrac{4}{\varepsilon+2h_M}e^{-2\delta(\varepsilon+h_M)}Y_0^T(t)Q_0Y_0(t),
\end{aligned} \tag{39}$$

where we utilize

$$\begin{aligned}
&2\int_0^1(\varepsilon\zeta + h_M)N_1^T(t - \varepsilon\zeta)Q_0N_1(t - \varepsilon\zeta)d\zeta \\
&= 2\int_0^1(\varepsilon\zeta + h_M)\mathrm{sq}^2\left(\tfrac{(t-\varepsilon\zeta-\eta)l_1}{\varepsilon}\right)\mathrm{sq}^2\left(\tfrac{(t-\varepsilon\zeta-\eta)l_2}{\varepsilon}\right)Q_0d\zeta \\
&= (\varepsilon + 2h_M)Q_0,
\end{aligned} \tag{40}$$

and the extended Jensen's inequality

$$\begin{aligned}
&2Y_0^T(t)Q_0Y_0(t) = 2\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^{t}\dot{\tilde{\theta}}^T(s)N_1^T(t - \varepsilon\zeta)dsd\zeta \\
&\quad \times Q_0\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^{t}N_1(t - \varepsilon\zeta)\dot{\tilde{\theta}}(s)dsd\zeta \\
&\leq 2\int_0^1(\varepsilon\zeta + h(t - \varepsilon\zeta))d\zeta\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^{t} \\
&\quad \times\dot{\tilde{\theta}}^T(s)N_1^T(t - \varepsilon\zeta)Q_0N_1(t - \varepsilon\zeta)\dot{\tilde{\theta}}(s)dsd\zeta \\
&< 2\int_0^1(\varepsilon\zeta + h_M)d\zeta\int_0^1\int_{t-\varepsilon\zeta-h_M}^{t} \\
&\quad \times\dot{\tilde{\theta}}^T(s)N_1^T(t - \varepsilon\zeta)Q_0N_1(t - \varepsilon\zeta)\dot{\tilde{\theta}}(s)dsd\zeta \\
&= (\varepsilon + 2h_M)\int_0^1\int_{t-\varepsilon\zeta-h_M}^{t} \\
&\quad \times\dot{\tilde{\theta}}^T(s)N_1^T(t - \varepsilon\zeta)Q_0N_1(t - \varepsilon\zeta)\dot{\tilde{\theta}}(s)dsd\zeta.
\end{aligned} \tag{41}$$

To compensate $Y_2(t)$ in (31), we define

$$\begin{aligned}
V_{Q_2}(t) &= q_2\int_0^1\int_{t-\varepsilon\zeta-h_M}^{t} e^{-2\delta(t-s)}(s - t + \varepsilon\zeta + h_M) \\
&\quad \times\dot{\tilde{\theta}}^T(s)H^T\tilde{\theta}(s)N_2^T(t - \varepsilon\zeta)N_2(t - \varepsilon\zeta)\tilde{\theta}^T(s)H\dot{\tilde{\theta}}(s) \\
&\quad \times dsd\zeta.
\end{aligned} \tag{42}$$

In this sense, we have

$$\begin{aligned}
&\dot{V}_{Q_2}(t)+2\delta V_{Q_2}(t)=q_2\dot{\tilde{\theta}}^T(t)H^T\tilde{\theta}(t)\textstyle\int_0^1(\varepsilon\zeta+h_M)N_2^T(t-\varepsilon\zeta)\\
&\quad\times N_2(t-\varepsilon\zeta)d\zeta\tilde{\theta}^T(t)H\dot{\tilde{\theta}}(t)\\
&\quad-q_2\textstyle\int_0^1\int_{t-\varepsilon\zeta-h_M}^t e^{-2\delta(t-s)}\dot{\tilde{\theta}}^T(s)H^T\tilde{\theta}(s)N_2^T(t-\varepsilon\zeta)\\
&\quad\times N_2(t-\varepsilon\zeta)\tilde{\theta}^T(s)H\dot{\tilde{\theta}}(s)dsd\zeta\\
&\le(\varepsilon+2h_M)q_2\dot{\tilde{\theta}}^T(t)H^T\tilde{\theta}(t)\tilde{\theta}^T(t)H\dot{\tilde{\theta}}(t)\\
&\quad-e^{-2\delta(\varepsilon+h_M)}q_2\textstyle\int_0^1\int_{t-\varepsilon\zeta-h_M}^t\dot{\tilde{\theta}}^T(s)H^T\tilde{\theta}(s)N_2^T(t-\varepsilon\zeta)\\
&\quad\times N_2(t-\varepsilon\zeta)\tilde{\theta}^T(s)H\dot{\tilde{\theta}}(s)dsd\zeta\\
&<(\varepsilon+2h_M)q_2\dot{\tilde{\theta}}^T(t)H^T\tilde{\theta}(t)\tilde{\theta}^T(t)H\dot{\tilde{\theta}}(t)\\
&\quad-\tfrac{2q_2}{\varepsilon+2h_M}e^{-2\delta(\varepsilon+h_M)}Y_2^T(t)Y_2(t),
\end{aligned}\tag{43}$$

where we have employed

$$\begin{aligned}
&\textstyle\int_0^1(\varepsilon\zeta+h_M)N_2^T(t-\varepsilon\zeta)N_2(t-\varepsilon\zeta)d\zeta=\\
&\textstyle\int_0^1(\varepsilon\zeta+h_M)\left[\mathrm{sq}\left(\frac{(t-\varepsilon\zeta-\eta)l_1}{\varepsilon}\right)\ \mathrm{sq}\left(\frac{(t-\varepsilon\zeta-\eta)l_2}{\varepsilon}\right)\right]\begin{bmatrix}\mathrm{sq}\left(\frac{(t-\varepsilon\zeta-\eta)l_1}{\varepsilon}\right)\\ \mathrm{sq}\left(\frac{(t-\varepsilon\zeta-\eta)l_2}{\varepsilon}\right)\end{bmatrix}\\
&\quad\times d\zeta=\varepsilon+2h_M,
\end{aligned}\tag{44}$$

and the extended Jensen's inequality

$$\begin{aligned}
&2q_2Y_2^T(t)Y_2(t)=2q_2\textstyle\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^t\dot{\tilde{\theta}}^T(s)H^T\tilde{\theta}(s)\\
&\quad\times N_2^T(t-\varepsilon\zeta)dsd\zeta\textstyle\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^t N_2(t-\varepsilon\zeta)\\
&\quad\times\tilde{\theta}^T(s)H\dot{\tilde{\theta}}(s)dsd\zeta\\
&\le 2q_2\textstyle\int_0^1(\varepsilon\zeta+h(t-\varepsilon\zeta))d\zeta\cdot\int_0^1\int_{t-\varepsilon\zeta-h(t-\varepsilon\zeta)}^t\dot{\tilde{\theta}}^T(s)H^T\tilde{\theta}(s)\\
&\quad\times N_2^T(t-\varepsilon\zeta)N_2(t-\varepsilon\zeta)\tilde{\theta}^T(s)H\dot{\tilde{\theta}}(s)dsd\zeta\\
&<2q_2\textstyle\int_0^1(\varepsilon\zeta+h_M)d\zeta\cdot\int_0^1\int_{t-\varepsilon\zeta-h_M}^t\dot{\tilde{\theta}}^T(s)H^T\tilde{\theta}(s)\\
&\quad\times N_2^T(t-\varepsilon\zeta)N_2(t-\varepsilon\zeta)\tilde{\theta}^T(s)H\dot{\tilde{\theta}}(s)dsd\zeta\\
&<(\varepsilon+2h_M)q_2\textstyle\int_0^1\int_{t-\varepsilon\zeta-h_M}^t\dot{\tilde{\theta}}^T(s)H^T\tilde{\theta}(s)\\
&\quad\times N_2^T(t-\varepsilon\zeta)N_2(t-\varepsilon\zeta)\tilde{\theta}^T(s)H\dot{\tilde{\theta}}(s)dsd\zeta.
\end{aligned}\tag{45}$$

Combining (8) with (11)-(15) and considering (24), we rewrite $\dot{\tilde{\theta}}(t)$ as

$$\begin{aligned}
\dot{\tilde{\theta}}(t)&=KH\left[\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right]+KH\tilde{\theta}(t-\eta)\\
&\quad+\bar{K}HN_1(t)\left[\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right]+\bar{K}HN_1(t)\tilde{\theta}(t-\eta)\\
&\quad+\tfrac{\bar{\bar{K}}}{2}N_2(t)\left[\tilde{\theta}^T(s_k)H\tilde{\theta}(s_k)-\tilde{\theta}^T(t-\eta)H\tilde{\theta}(t-\eta)\right]\\
&\quad+\tfrac{\bar{\bar{K}}}{2}N_2(t)\tilde{\theta}^T(t-\eta)H\tilde{\theta}(t-\eta)+\omega(t),\quad t\in[t_k,t_{k+1}).
\end{aligned}\tag{46}$$

To compensate the error term $\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)$ and the state with constant delay $\tilde{\theta}(t-\eta)$ in (46), we employ

$$\begin{aligned}
&V_{S_0}(t)=\textstyle\int_{t-\eta}^t e^{-2\delta(t-s)}\tilde{\theta}^T(s)S_0\tilde{\theta}(s)ds,\\
&V_{R_0}(t)=\eta\textstyle\int_{t-\eta}^t e^{-2\delta(t-s)}(s-t+\eta)\dot{\tilde{\theta}}^T(s)R_0\dot{\tilde{\theta}}(s)ds,\\
&V_W(t)=(h_M-\eta)^2e^{2\delta(h_M-\eta)}\textstyle\int_{s_k}^t e^{-2\delta(t-s)}\dot{\tilde{\theta}}^T(s)W\dot{\tilde{\theta}}(s)ds\\
&\quad-\tfrac{\pi^2}{4}\textstyle\int_{s_k}^{t-\eta}e^{-2\delta(t-s)}\left[\tilde{\theta}(s_k)-\tilde{\theta}(s)\right]^TW\left[\tilde{\theta}(s_k)-\tilde{\theta}(s)\right]ds,\\
&\quad t\in[t_k,t_{k+1}),
\end{aligned}\tag{47}$$

where $V_W(t)\ge 0$ by the extended Wirtinger's inequality, and it does not grow in the jumps $t=s_k$. Thus, we have

$$\begin{aligned}
&\dot{V}_{S_0}(t)+2\delta V_{S_0}(t)=\tilde{\theta}^T(t)S_0\tilde{\theta}(t)\\
&\quad-e^{-2\delta\eta}\tilde{\theta}^T(t-\eta)S_0\tilde{\theta}(t-\eta),\\
&\dot{V}_{R_0}(t)+2\delta V_{R_0}(t)\le\eta^2\dot{\tilde{\theta}}^T(t)R_0\dot{\tilde{\theta}}(t)\\
&\quad-e^{-2\delta\eta}\left[\tilde{\theta}(t)-\tilde{\theta}(t-\eta)\right]^TR_0\left[\tilde{\theta}(t)-\tilde{\theta}(t-\eta)\right],\\
&\dot{V}_W(t)+2\delta V_W(t)=\tfrac{\varepsilon^2}{16}e^{\frac{\delta\varepsilon}{2}}\dot{\tilde{\theta}}^T(t)W\dot{\tilde{\theta}}(t)-\tfrac{\pi^2}{4}e^{-2\delta\eta}\\
&\quad\times\left[\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right]^TW\left[\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right],t\in[t_k,t_{k+1}).
\end{aligned}\tag{48}$$

Define the Lyapunov-Krasovskii functional (LKF) as

$$\begin{aligned}
V(t)&=V_P(t)+V_R(t)+V_{Q_1}(t)+V_{Q_0}(t)+V_{Q_2}(t)\\
&\quad+V_{S_0}(t)+V_{R_0}(t)+V_W(t).
\end{aligned}\tag{49}$$

and assume the overall bound

$$\left|\tilde{\theta}(t)\right|<\sigma,\quad t\ge 0,\tag{50}$$

to address the nonlinear terms. Taking into account that

$$\left[N_1(t)\tilde{\theta}(t-\eta)\right]^2\le\tilde{\theta}^2(t-\eta),\tag{51}$$

$$\left[N_1(t)\left(\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right)\right]^2\le\left[\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right]^2,\tag{52}$$

$$\begin{aligned}
&N_2^2(t)\left[\tilde{\theta}^T(t-\eta)H\tilde{\theta}^T(t-\eta)\right]^2\\
&\le 2\left[\lambda(H)\tilde{\theta}^2(t-\eta)\right]^2<2\lambda^2(H)\sigma^2\tilde{\theta}^2(t-\eta),
\end{aligned}\tag{53}$$

$$\begin{aligned}
&N_2^2(t)\left[\tilde{\theta}^T(s_k)H\tilde{\theta}(s_k)-\tilde{\theta}^T(t-\eta)H\tilde{\theta}(t-\eta)\right]^2\\
&\le 2\lambda_{\max}(H^2)\left(\tilde{\theta}(s_k)+\tilde{\theta}(t-\eta)\right)^2\left(\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right)^2\\
&<8\lambda_{\max}(H^2)\sigma^2\left[\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right]^2,
\end{aligned}\tag{54}$$

Considering (31), (33), (36), (39), (43), (48), (51)-(54), we have

$$\begin{aligned}
&\dot{V}(t)+2\delta V(t)+\lambda_1\left[\tilde{\theta}^2(t-\eta)-\left[N_1(t)\tilde{\theta}(t-\eta)\right]^2\right]\\
&\quad+\lambda_2\left[\left[\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right]^2-\left[N_1(t)\left(\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right)\right]^2\right]\\
&\quad+\lambda_3\Big[2\lambda^2(H)\sigma^2\tilde{\theta}^2(t-\eta)\\
&\quad-N_2^2(t)\left[\tilde{\theta}^T(t-\eta)H\tilde{\theta}(t-\eta)\right]^2\Big]-(\varepsilon+\eta)b\omega^2(t)\\
&\quad+\lambda_4\Big[8\lambda_{\max}(H^2)\sigma^2\left[\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right]^2\\
&\quad-N_2^2(t)\left[\tilde{\theta}^T(s_k)H\tilde{\theta}(s_k)-\tilde{\theta}^T(t-\eta)H\tilde{\theta}(t-\eta)\right]^2\Big]\\
&\quad-(\varepsilon+\eta)b\omega^2(t)\\
&<\xi^T(t)\Omega\xi(t)+\dot{\tilde{\theta}}^T(t)\left(\varepsilon\Xi_\varepsilon+\eta\Xi_\eta\right)\dot{\tilde{\theta}}(t)\\
&=\xi^T(t)\Omega\xi(t)+\xi^T(t)\left[\Psi_\varepsilon\ \Psi_\eta\right]\begin{bmatrix}\varepsilon{\Xi_\varepsilon}^{-1}&0\\0&\eta{\Xi_\eta}^{-1}\end{bmatrix}\begin{bmatrix}\Psi_\varepsilon^T\\ \Psi_\eta^T\end{bmatrix}\xi(t)<0,
\end{aligned}\tag{55}$$

where $\xi(t)=\Big[\tilde{\theta}(t),G(t),Y_1(t),Y_0(t),Y_2(t),\tilde{\theta}(t-\eta),\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta),N_1(t)\tilde{\theta}(t-\eta),N_1(t)\left(\tilde{\theta}(s_k)-\tilde{\theta}(t-\eta)\right),N_2(t)\tilde{\theta}^T(t-\eta)H\tilde{\theta}(t-\eta),N_2(t)\left[\tilde{\theta}^T(s_k)H\tilde{\theta}(s_k)-\tilde{\theta}^T(t-\eta)H\tilde{\theta}(t-\eta)\right],\omega(t)\Big]^T$, and we utilize $\tilde{\theta}(t)\tilde{\theta}^T(t)\le\sigma^2I$. The inequality (55) follows from $\Phi_2<0$ in (25) by Schur complement.

• Step 2—we prove the assumed overall bound (50) is not violated. In implementation, usually $\dot{\tilde{\theta}}(t)=\dot{\hat{\theta}}(t)\equiv 0$ when $t<0$. It means that the update law $\dot{\hat{\theta}}(t)$ is turned off and the estimate $\hat{\theta}(t)$ is fixed at some constant which renders the initial condition of the error

$$\left|\tilde{\theta}(t)\right|=\left|\hat{\theta}-\theta^*\right|=\left|\tilde{\theta}(0)\right|\le\sigma_0<\sigma,\quad t<0.\tag{56}$$

Considering (11), for $\tau \in [-h_M, 0]$,

$$\tilde{\theta}(t+\tau) = \begin{cases} \tilde{\theta}(0), \quad t+\tau < 0, \\ \tilde{\theta}(0) + \int_0^{t+\tau} KM(s-\eta)\Big[S^T(s-\eta)H \\ \quad \times \tilde{\theta}(s-h(s)) + \frac{1}{2}\tilde{\theta}^T(s-h(s))H\tilde{\theta}(s-h(s)) \\ \quad +\omega(s)\Big]ds, \quad t+\tau \geq 0. \end{cases} \tag{57}$$

where $\tau \in [-h_M, 0]$. Then, we arrive at

$$\left\|\tilde{\theta}_t\right\|_C \leq \left|\tilde{\theta}(0)\right| + (\varepsilon+\eta)\bar{\omega} + \int_0^t \mu \left\|\tilde{\theta}_s\right\|_C ds, \quad t \in [0, \varepsilon+\eta], \tag{58}$$

where $\bar{\omega}$ is defined by (13) and μ is given in (27). By Gronwall inequality, the following inequality follows from (58)

$$\left|\tilde{\theta}(t)\right| \leq \left\|\tilde{\theta}_t\right\|_C < \left(\left|\tilde{\theta}(0)\right| + (\varepsilon+\eta)\bar{\omega}\right)e^{\mu t} < \sigma, \quad t \in [0, \varepsilon+\eta]. \tag{59}$$

The latter corresponds to the first formula in (28) and is ensured by $\Phi_3 < \sigma^2$ in (25). The equations (8)-(10) result in

$$\left|\dot{\tilde{\theta}}(t)\right| < \sqrt{\frac{k_1^2}{a_1^2} + \frac{k_2^2}{a_2^2}}\left[|Q^*| + \frac{\lambda(H)}{2}\left((\sigma_0 + (\varepsilon+\eta)\bar{\omega})e^{\mu(\varepsilon+\eta)} + \sqrt{a_1^2 + a_2^2}\right)^2\right] = \Delta, \quad t \in [0, \varepsilon+\eta]. \tag{60}$$

With Jensen's inequality, we obtain

$$V(t) \geq V_P(t) + V_R(t) \geq \begin{bmatrix} \tilde{\theta}(t) \\ G(t) \end{bmatrix}^T \begin{bmatrix} P & -P \\ * & P + e^{-2\delta\varepsilon}R \end{bmatrix} \begin{bmatrix} \tilde{\theta}(t) \\ G(t) \end{bmatrix} \geq \left|\tilde{\theta}(t)\right|^2, \tag{61}$$

which follows from $\Phi_1 > 0$ in (25).

From (55), we have

$$\dot{V}(t) + 2\delta V(t) < (\varepsilon+\eta)b\omega^2(t) \leq (\varepsilon+\eta)b\bar{\omega}^2, \tag{62}$$

where $\bar{\omega}$ is given by (13).

Applying comparison principle to (62), we have

$$V(t) < V(\varepsilon+\eta)e^{-2\delta(t-\varepsilon-\eta)} + \left(1 - e^{-2\delta(t-\varepsilon-\eta)}\right)\frac{\varepsilon+\eta}{2\delta}b\bar{\omega}^2, \quad t \geq \varepsilon+\eta. \tag{63}$$

From (30), we have

$$V_P(\varepsilon+\eta) < \left(1 + \frac{1}{q}\right)\lambda_P\left[\left(\left|\tilde{\theta}(0)\right| + (\varepsilon+\eta)\bar{\omega}\right)^2 e^{2\mu(\varepsilon+\eta)}\right] + \frac{\varepsilon^2(1+q)}{4}\lambda_P\Delta^2. \tag{64}$$

From (32), we get

$$V_R(\varepsilon+\eta) = \frac{1}{\varepsilon}\int_\eta^{\varepsilon+\eta} e^{-2\delta(\varepsilon+\eta-\tau)}(\tau-\eta)^2\dot{\tilde{\theta}}^T(\tau)R\dot{\tilde{\theta}}(\tau)d\tau < \frac{\lambda_R}{\varepsilon}\int_\eta^{\varepsilon+\eta}(\tau-\eta)^2 d\tau \cdot \Delta^2 = \frac{\varepsilon^2}{3}\lambda_R\Delta^2. \tag{65}$$

From (35), we have

$$V_{Q_1}(\varepsilon+\eta) < \frac{1}{3}(\varepsilon^2 + 3\varepsilon h_M + 3h_M^2)\lambda_{Q_1}\Delta^2. \tag{66}$$

From (38), we obtain

$$V_{Q_0}(\varepsilon+\eta) = 2\int_0^1 \int_{\varepsilon+\eta-\varepsilon\zeta-h_M}^{\varepsilon+\eta} e^{-2\delta(\varepsilon+\eta-s)} < \frac{1}{3}(\varepsilon^2 + 3\varepsilon h_M + 3h_M^2)\lambda_{Q_0}\Delta^2. \tag{67}$$

From (42), we get

$$V_{Q_2}(\varepsilon+\eta) < \frac{q_2}{3}(\varepsilon^2 + 3\varepsilon h_M + 3h_M^2)\lambda_{\max}(H^2)\sigma^2\Delta^2. \tag{68}$$

From (47), we have

$$\begin{aligned} &V_{S_0}(\varepsilon+\eta) = \int_\varepsilon^{\varepsilon+\eta} e^{-2\delta(\varepsilon+\eta-s)}\tilde{\theta}^T(s)S_0\tilde{\theta}(s)ds < \eta\lambda_{S_0}\sigma^2, \\ &V_{R_0}(\varepsilon+\eta) = \eta\int_\varepsilon^{\varepsilon+\eta} e^{-2\delta(\varepsilon+\eta-s)}(s-\varepsilon)\dot{\tilde{\theta}}^T(s)R_0\dot{\tilde{\theta}}(s)ds \\ &< \frac{1}{2}\eta^3\lambda_{R_0}\Delta^2, \\ &V_W(\varepsilon+\eta) \leq (h_M-\eta)^2 e^{2\delta(h_M-\eta)}\int_{s_k}^{t_{k+1}}\dot{\tilde{\theta}}^T(s)W\dot{\tilde{\theta}}(s)ds \\ &< \frac{h_M\varepsilon^2}{16}e^{\frac{\delta\varepsilon}{2}}\lambda_W\Delta^2. \end{aligned} \tag{69}$$

Plugging (64)-(69) into (63), we arrive at the second formula in (28) which follows from $\Phi_4 < \sigma^2$ in (25). Finally, by contradiction-based arguments of [20, Appendix A] it can be proved that (25) implies (50).

References

[1] Ariyur, K. B., Krstic, M. (2003), *Real-time optimization by extremum-seeking control*, Wiley.

[2] Durr, H.-B., Stankovic, M. S., Ebenbauer, C., Johansson, K. H. (2013), Lie bracket approximation of extremum seeking systems, *Automatica*, vol. 49, pp. 1538-1552.

[3] Fridman, E. (2014), *Introduction to Time-Delay Systems*, Springer.

[4] Fridman, E., Zhang, J. (2020), Averaging of linear systems with almost periodic coefficients: a time-delay aproach, *Automatica*, vol. 122, 109287.

[5] Fu, W., Qin, J., Zheng, W. X., Chen, Y., Kang, Yu., (2021), Resilient cooperative source seeking of double-integrator multi-robot systems under deception attacks, *IEEE Transactions on Industrial Electronics*, vol. 68, pp. 4218-4227.

[6] Galarza-Jimenez, F., Poveda, J. I., Bianchin, G., Anese, E. D., (2021), Extremum seeking under persistent gradient deception: a switching systems approach, *IEEE Control Systems Letters*, vol. 6, pp. 133-138.

[7] Hazeleger, L., Beerens, R., Wouw, N. V. D. (2019), A sampled-data extremum-seeking approach for accurate setpoint control of motion systems with friction, *IFAC-PapersOnLine*, vol. 52, pp. 801-806.

[8] Khalil, H. K. (2002). *Nonlinear systems*, Prentice Hall.

[9] Krstic, M., Wang H.-H. (2000), Stability of extremum seeking feedback for general nonlinear dynamic systems, *Automatica*, vol. 36, pp. 595-601.

[10] Khong S. Z., Nesic, D., Tan, Y., Manzie, C., (2013), Unified frameworks for sampled-data extremum seeking control: global optimisation and multi-unit systems, *Automatica*, vol. 49, pp. 2720-2733.

[11] Oliveira, T. R., Krstic, M., Tsubakino, D. (2016), Extremum seeking for static maps with delays, *IEEE Transactions on Automatic Control*, vol. 62, pp. 1911-1926.

[12] Oliveira, T. R., Feiling, J., Koga, S., Krstic, M. (2020), Multivariable extremum seeking for PDE dynamic systems, *IEEE Transactions on Automatic Control*, vol. 65, pp. 4949-4956.

[13] Poveda, J. I., Teel, A. R. (2017), A framework for a class of hybrid extremum seeking controllers with dynamic inclusions, *Automatica*, vol. 76, pp. 113-126.

[14] Poveda, J. I., Teel, A. R. (2017), A robust event-triggered approach for fast sampled-data extremization and learning, *IEEE Transactions on Automatic Control*, vol. 62, pp. 4949-4964.

[15] Premaratne, U., Halgamuge, S., Tan, Y., Mareels, I. M. Y. (2020), Extremum seeking control with sporadic packet transmission for networked control systems, *IEEE Transactions on Control of Network Systems*, vol. 7, pp. 758-769.

[16] Scheinker, A., Krstic, M. (2017), *Model-free stabilization by extremum seeking*, Springer.

[17] Scheinker, A., Scheinker, D., (2016), Bounded extremum seeking with discontinuous dithers, *Automatica*, vol. 69, pp. 250-257.

[18] Tan, Y., Nesic, D., Mareels, I., (2008), On the choice of dither in extremum seeking systems: a case study, *Automatica*, vol. 44, pp. 1446-1450.

[19] Teel, A. R., Popovic, D., (2001), Solving smooth and nonsmooth multivariable extremum seeking problems by the methods of nonlinear programming, Proceedings of the 2001 American Control Conference, pp. 2394-2399, Arlington, VA, USA, June 25-27, 2001.

[20] Zhu, Y., Fridman, E. (2022), Extremum seeking via a time-delay approach to averaging, *Automatica*, vol. 135, 109965.

[21] Zhu, Y., Fridman, E., Oliveira, T. R. (2021), A time-delay approach for sampled-data and delayed scalar Extremum seeking, *2021 IEEE Conference on Decision and Control*, accepted.