Automatica 166 (2024) 111710

Contents lists available at ScienceDirect

Automatica

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

Extremum seeking of general nonlinear static maps: A time-delay approach☆

Gaofeng Pan [a,b], Yang Zhu [a,b,*], Emilia Fridman [c], Zhengguang Wu [a]

[a] College of Control Science and Engineering, State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China
[b] Ningbo Innovation Center, Zhejiang University, Ningbo, China
[c] School of Electrical Engineering, Tel-Aviv University, Israel

ARTICLE INFO

Article history:
Received 3 April 2023
Received in revised form 1 March 2024
Accepted 15 April 2024
Available online xxxx

Keywords:
Extremum seeking
Nonlinear
Static
Time-delay

ABSTRACT

In this paper, we extend a newly developed time-delay approach of extremum seeking (ES) from quadratic maps to general nonlinear maps, which substantially expands the scope of application of the time-delay method. The gradient-based ES in the case of single-variable and multi-variable nonlinear static maps are considered. Different from the recent literature, the time-delay method in this paper is applicable to a big family of nonlinear maps satisfying a couple of mild assumptions. By transforming the original ES system into a kind of time-delay system of neutral type, and further transforming the neutral delay differential equation into the perturbed ordinary differential equation (ODE), we provide simple inequalities that guarantee practical stability of the ES closed-loop system. With a prior knowledge about the upper bounds of the nonlinear map and its gradient and Hessian, the time-delay approach suggests a quantitative calculation on the lower bound of dither frequency and the ultimate upper bound of estimation error, which is difficult to achieve by the classical averaging method.

1. Introduction

Extremum Seeking (ES) is an adaptive control strategy to search for optimum point of system operation. The idea of ES witnessed its first appearance in 1922 (Tan, Moase, Manzie, Nešić, & Mareels, 2010). Since Krstic & Wang gave the first rigorous proof for the algorithm's convergence in 2000 (Krstić & Wang, 2000), ES has been kept as an active field of research in control theory. Various ES improvements and applications can be found in the recent literature: semi-global stability (Tan, Nešić, & Mareels, 2006), Lie-bracket approximation (Dürr, Stanković, Ebenbauer, & Johansson, 2013; Labar, Ebenbauer, & Marconi, 2022), time-varying ES (Guay & Dochain, 2015) and uncertainty estimation in ES (Guay, 2021), sampled-data ES (Hazeleger, Nesic, & van de Wouw, 2022; Zhu, Fridman, & Oliveira, 2023), ES for multi-agent systems (Haring, Fossy, Silva, & Pavlov, 2022; Krilasevic & Grammatico, 2021), ES with stochastic averaging (Liu & Krstic, 2012), ES with unknown control direction (Mele, Tommasi, & Pironti, 2022; Scheinker & Krstić, 2017), ES with time-delay (Malisoff & Krstic, 2021; Oliveira, Krstic, & Tsubakino, 2016; Zhu & Krstic, 2020), etc.

It is well-known that ES control systems are practically stable provided tuning parameters meet some conditions like the dither frequency should be large whereas the dither amplitude should be small (Ariyur & Krstic, 2003). However, none of existing asymptotic methods suggest quantitative bounds on these tuning parameters that preserve the stability. Having an analytically known bound for the choice of tuning parameters is of importance for practical applications for which a proper bound on tuning parameters must typically be found by trial and error.

Recently a novel time-delay approach to periodic averaging has been presented in Fridman and Zhang (2020) in which efficient bounds on the small parameter that preserves the system stability are offered. Later on, the papers Zhu and Fridman (2022), Zhu et al. (2023) proposed a constructive time-delay approach for ES in the case of both continuous and sampled-data control. See also Zhang and Fridman (2023) for Lie-brackets-based averaging of affine systems via a time-delay approach. Different from the classical averaging method, the time-delay approach in Zhu and Fridman (2022) and Zhu et al. (2023) does not use any approximation, and for the first time gives quantitative bounds on the tuning parameters and extremum seeking error by solving linear matrix inequalities (LMIs).

Unfortunately, the time-delay approach to ES developed in Zhu and Fridman (2022) and Zhu et al. (2023) is limited to static maps

☆ A preliminary version for scalar maps was presented on IFAC NOLCOS 2023. This paper was recommended for publication in revised form by Associate Editor Raul Ordonez under the direction of Editor Miroslav Krstic.

* Corresponding author at: College of Control Science and Engineering, State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China, and Ningbo Innovation Center, Zhejiang University, Ningbo, China.

E-mail addresses: 12132047@zju.edu.cn (G. Pan), zhuyang88@zju.edu.cn (Y. Zhu), emilia@tauex.tau.ac.il (E. Fridman), nashwzhg@zju.edu.cn (Z. Wu).

https://doi.org/10.1016/j.automatica.2024.111710
0005-1098/

of a standard quadratic form, which cannot capture complex and diverse dynamics in practice. In this paper, we remove this "quadratic" restriction so that the time-delay approach is available for ES of general nonlinear static maps, which substantially expands the application range of the method. Motivated by Yang and Fridman (2023) where the map considered is still confined to a quadratic form, instead of analyzing the time-delay system of neutral type directly, we further transform the neutral delay differential equation into the retarded system with the nominal part in the form of the averaged ordinary differential equation (ODE), which renders the Lyapunov-based analysis rather simple and reduces the conservativeness of the bounds on tuning parameters and estimation error. A preliminary conference result for scalar systems via a more complicated analysis that employs Lyapunov–Krasovskii method for neutral type systems had been presented in Zhu and Fridman (2023).

The rest organization of the paper is as follows: In Section 2, we apply the time-delay approach to the gradient-based ES of scalar static maps. Section 3 extends the time-delay approach to vector static maps. Section 4 provides two examples with simulation results, and Section 5 summarizes some conclusions.

Notation:

- For a scalar $x \in \mathbb{R}$, we denote $|x|$ to be the absolute value of x.
- For a vector $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$, we denote $|x|$ to be the Euclidean norm $|x| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$.
- For a matrix $A \in \mathbb{R}^{n\times n}$, we denote $|A|$ to be the 2-norm $|A| = \sqrt{\lambda_{\max}\left(A^T A\right)}$.
- For a function $q(\theta) : D \to \mathbb{R}$, by $q(\theta) \in C^2(D)$, we refer to that $q(\theta)$ has continuous derivatives up to the 2nd-order in the domain $\theta \in D$.

2. Scalar systems

For conceptional clearness, we start with scalar systems in this section. The more general vector systems will be discussed in the next section.

Consider single-variable nonlinear static maps

$$y(t) = q(\theta(t)) \tag{1}$$

where $y(t) \in \mathbb{R}$ is the measurable output and $\theta(t) \in \mathbb{R}$ is the control input. The map (1) satisfies the following assumptions.

Assumption 1. There exist $\theta^* \in \mathbb{R}$, $\sigma > 0$ and small $a > 0$ such that $q(\theta) \in C^2(\theta^* - \sigma - a, \theta^* + \sigma + a)$, and the following hold:

$$q'\left(\theta^*\right) = 0, \quad q''(\theta^*) = H < 0, \tag{2}$$

$$q'\left(\theta^* + \Delta\right) \cdot \Delta \le -\mu(\sigma) \cdot \Delta^2 < 0, \quad \forall 0 < |\Delta| < \sigma \tag{3}$$

where $\mu(\sigma) > 0$ is a known σ-dependent constant, which decreases monotonically in σ.

In Assumption 1, the condition (3) is not restrictive:

$$\begin{aligned} q'(\theta^* + \Delta) \cdot \Delta &= \left[q'(\theta^*) + \int_{\theta^*}^{\theta^*+\Delta} q''(s)\, ds\right] \cdot \Delta \\ &\overset{s=\theta^*+\xi\Delta}{=\!=\!=} \int_0^1 q''(\theta^* + \xi\Delta)\, d\xi \cdot \Delta^2 \end{aligned} \tag{4}$$

The inequality $q''(\theta^*) < 0$ and the continuity of $q''(\theta)$ indicates that there always exists $\sigma > 0$ such that for all deviations $|\Delta| < \sigma$, the average from $q''(\theta^*)$ to $q''(\theta^* + \Delta)$ is negative. Thus,

$$\int_0^1 q''(\theta^* + \xi\Delta)\, d\xi \le -\mu(\sigma) < 0, \quad \forall 0 < |\Delta| < \sigma. \tag{5}$$

Then, (3) holds.

For Assumption 1, note that:

- Eq. (2) suggests the existence of the extremum. Namely, $\theta = \theta^*$ is an input extremum of the map $q(\theta)$ at which the output has a local maximum $y = q(\theta^*)$. Without loss of generality, the maximum case is taken into account. The minimum case could be handled in a parallel way under the assumption with the opposite sign such that $q'(\theta^*) = 0$, $q''(\theta^*) > 0$, $q'(\theta^* + \Delta) \cdot \Delta \ge \mu(\sigma) \cdot \Delta^2 > 0$.
- Condition (3) indicates the uniqueness of the extremum such that $q'(\theta) \neq 0$, $\forall \theta \neq \theta^*$, $\theta \in (\theta^* - \sigma - a, \theta^* + \sigma + a)$. The extremum uniqueness within the σ-neighborhood is important since such a neighborhood infers the region of attraction of ES (see Example 1), which is consistent with the local stability of ES in the literature. For practical implementation of ES algorithm, it is necessary to determine σ-value of an unknown map in advance by some techniques like system identification, empirical method, and trail and error.

Assumption 2. For any $|\Delta| < \sigma$ and a defined in Assumption 1, given $\xi \in [-1, 1]$, we have

$$\begin{aligned} &|q(\theta^* + \Delta)| < q_0(\sigma), \\ &\left|q'(\theta^* + \Delta)\right| < q_1(\sigma), \quad \left|q''(\theta^* + \Delta)\right| < q_2(\sigma), \\ &\left|q''(\theta^* + \Delta + a\xi) - q''(\theta^*)\right| < L\,|\Delta + a\xi| \end{aligned} \tag{6}$$

where $q_0(\sigma)$, $q_1(\sigma)$, $q_2(\sigma)$, L are known positive constants.

In the literature, the map (1) is usually unknown, the bounds on the map and its derivatives are given for analysis but are not supposed to be known (Guay & Dochain, 2015; Tan et al., 2006). In the face of a "black box" model, it is hard to choose tuning parameters for a reliable control and even to perform simulations.

Here we study a "grey box" by Assumption 2: the map $q(\theta)$ is unknown, but a few of bounds of the map and its derivatives (up to the 2nd-order) are known, and the 2nd-order derivative is local Lipschitz with a known Lipschitz constant (here the Lipschitz condition is not necessary, but it reduces the conservativeness of the LMI-based analysis later. See the bound (29) of $\mathcal{R}(t)$ defined in (17) and Remark 3). When the knowledge of the bounds in (6) is unavailable, the time-delay approach proposed in the paper presents a qualitative analysis in a more accurate way comparative to the classical averaging method (see Remarks 1–2). When the prior knowledge in (6) is available, the time-delay method allows a quantitative analysis by which we are able to calculate the upper bound of the key tuning parameter (the dither period ε given in (19)) and the ultimate bound of the estimation error of ES (defined by (10)). A compromise between the quantitative analysis with the plant information and the qualitative analysis without the model knowledge is always there. The more precisely we know the map bounds specified in Assumptions 1–2, the more accurately we are able to analyze the system's performance and estimate the tuning parameters' range.

Example 1. Consider the following nonlinear static map whose trajectory is shown in Fig. 1:

$$q(\theta) = \tfrac{1}{3}\theta^3 - \theta \tag{7}$$

It is seen that $\theta^* = -1$ and $\theta^* = 1$ are the local maximum and minimum, respectively. At the local maximum $\theta^* = -1$, the derivatives up to the 2nd-order are as follows:

$$\begin{cases} q'(\theta^*) = \theta^2 - 1\Big|_{\theta=\theta^*=-1} = 0 \\ q''(\theta^*) = 2\theta\Big|_{\theta=\theta^*=-1} = -2 < 0 \end{cases} \tag{8}$$

which means that the property (2) in Assumption 1 is satisfied. In order to meet the condition (3) in Assumption 1, we have

$$\begin{aligned} &q'(\theta^* + \Delta) \cdot \Delta = \left[(\theta^* + \Delta)^2 - 1\right]\Delta \\ &= \left[(\Delta - 1)^2 - 1\right]\Delta = (\Delta - 2)\,\Delta^2 \le -\mu(\sigma)\,\Delta^2 < 0 \end{aligned} \tag{9}$$

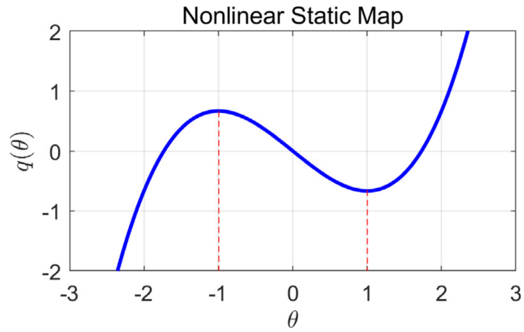


Fig. 1. Nonlinear static map: $q(\theta) = \frac{1}{3}\theta^3 - \theta$.

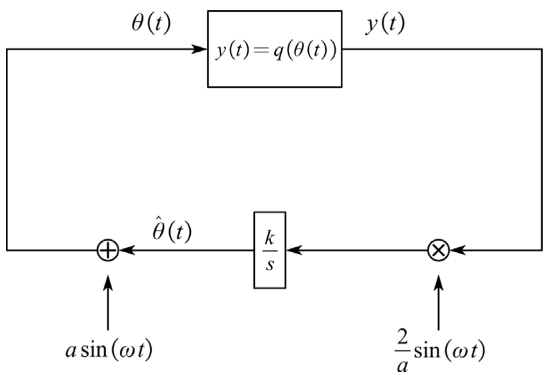


Fig. 2. ES for single-variable nonlinear static maps.

Thus $\Delta < \sigma < 2$, which indicates that the positive deviation to $\theta^* = -1$ cannot go beyond 2. In other words, $\theta^* + \Delta < 1$ for $\theta^* = -1$. This is consistent with the perceptual intuition from Fig. 1, which indicates that you cannot find the local maximum $\theta^* = -1$ if the initial value starts from any point such that $\theta(0) \geq 1$. ■

To render $\theta(t) \to \theta^*$, we introduce $\hat{\theta}(t)$ as a real time estimate of θ^* with the estimation error being

$$\tilde{\theta}(t) = \hat{\theta}(t) - \theta^* \tag{10}$$

We employ the gradient-based ES algorithm in Fig. 2,

$$\begin{aligned} &\theta(t) = \hat{\theta}(t) + a\sin(\omega t), \\ &\dot{\hat{\theta}}(t) = k \cdot \tfrac{2}{a}\sin(\omega t) \cdot y(t) \end{aligned} \tag{11}$$

where the sign of k is opposite to the sign of the Hessian $q''(\theta^*)$ and the initial value $\hat{\theta}(0) \in [\theta^* - \sigma_0, \theta^* + \sigma_0]$, where $\sigma_0 < \sigma$ is a known constant.

To analyze the convergence of ES (11), we consider Taylor expansion of (1) such that

$$\begin{aligned} &q(\theta(t)) = q\left(\theta^* + \tilde{\theta}(t) + a\sin(\omega t)\right) \\ &= q\left(\theta^* + \tilde{\theta}(t)\right) + q'\left(\theta^* + \tilde{\theta}(t)\right) a\sin(\omega t) \\ &\quad + \tfrac{q''\left(\theta^* + \tilde{\theta}(t) + \xi(t) a\sin(\omega t)\right)}{2} a^2 \sin^2(\omega t) \end{aligned} \tag{12}$$

where $\xi(t) \in (0, 1)$. An alternative Taylor formula is suggested as follows:

$$\begin{aligned} &q(\theta(t)) = q(\theta^*) + q'(\theta^*)\left(\tilde{\theta}(t) + a\sin(\omega t)\right) \\ &\quad + \tfrac{q''(\theta^*)}{2}\left(\tilde{\theta}(t) + a\sin(\omega t)\right)^2 \\ &\quad + \tfrac{q'''\left(\theta^* + \xi(t)(\tilde{\theta}(t) + a\sin(\omega t))\right)}{3!}\left(\tilde{\theta}(t) + a\sin(\omega t)\right)^3 \end{aligned}$$

where θ^* is treated as the baseline and $\tilde{\theta}(t) + a\sin(\omega t)$ is regarded as increment. Although the above equation is of a standard quadratic form plus a remainder term, such a handling indicates that we get one more restrictive assumption that the map $q(\theta)$ has the derivatives with respect to θ up to the 3rd-order. Furthermore, substituting the map $q(\theta(t))$ with the remainder $\frac{q'''(\cdot)}{3!}\left(\tilde{\theta}(t) + a\sin(\omega t)\right)^3$ into the update law (11), we will get the term $\frac{2k}{a}\sin(\omega t)\frac{q'''(\cdot)}{3!}\tilde{\theta}^3(t)$ in the error dynamical equation, which will be hard to address by the classical averaging or the time-delay method. This is due to the fact that $q'''(\cdot)$ depends upon the dither $\sin(\omega t)$ and is a fast time-varying variable whose average is not zero.

In the presence of σ defined in Assumption 1 and the definitions of (10)–(11), the overall bound of the estimation error is supposed to satisfy

$$\left|\tilde{\theta}(t)\right| < \sigma, \quad t \geq 0 \tag{13}$$

Then we have $\theta^* - \sigma - a < \theta(t) = \theta^* + \tilde{\theta}(t) + a\sin(\omega t) < \theta^* + \sigma + a$, so that the domain of θ is in line with the domain in Assumption 1. The bound (13) will be guaranteed by the LMI conditions (31) in Theorem 1. Notice that the domain defined by the overall bound (13) is not identical with the domain given by the ultimate bound (33) in Theorem 1 (see Fig. 4 and the explanation in the case of double-variable maps).

Substituting (12) into the 2nd equation of (11) and taking the time-derivative of (10), the dynamics of the estimation error is governed by

$$\begin{aligned} &\dot{\tilde{\theta}}(t) = \tfrac{2k}{a}\sin(\omega t) \cdot \Big[\, q\left(\theta^* + \tilde{\theta}(t)\right) + q'\left(\theta^* + \tilde{\theta}(t)\right) \\ &\quad \times a\sin(\omega t) + \tfrac{q''\left(\theta^* + \tilde{\theta}(t) + \xi(t) a\sin(\omega t)\right)}{2} a^2\sin^2(\omega t)\,\Big] \\ &= kq'\left(\theta^* + \tilde{\theta}(t)\right) - kq'\left(\theta^* + \tilde{\theta}(t)\right)\cos(2\omega t) \\ &\quad + \tfrac{2k}{a} q\left(\theta^* + \tilde{\theta}(t)\right)\sin(\omega t) + kaH(t)\sin^3(\omega t) \end{aligned} \tag{14}$$

where $H(t) = q''\left(\theta^* + \tilde{\theta}(t) + \xi(t) a\sin(\omega t)\right)$ is defined for notational simplicity, and $2\sin^2(\omega t) = 1 - \cos(2\omega t)$ is used.

Remark 1. To analyze the error system (14), a majority of existing literature on ES resort to averaging theory in Khalil (2002, Chapter 10.4). To be specific, by setting the dither frequency ω to be large while the adaptation gain k to be small, the state variable $\tilde{\theta}(t)$ is slowly time-varying in comparison with the dither signals $\cos(2\omega t)$, $\sin(\omega t)$, $\sin^3(\omega t)$ as if it is a "freezing" constant. Then, $q'\left(\theta^* + \tilde{\theta}(t)\right)$ and $q\left(\theta^* + \tilde{\theta}(t)\right)$ are treated as "freezing" constants as well. By selecting the dither amplitude a to be small, the last term $kaH(t)\sin^3(\omega t)$ on the right-hand side of (14) is small and thus neglected. Taking advantage of the fact that the average of $\sin(\omega t)$ and $\cos(2\omega t)$ over one period $T = \frac{2\pi}{\omega}$ are zeros, the averaged system is derived as

$$\begin{aligned} &\dot{\tilde{\theta}}_{av}(t) = kq'\left(\theta^* + \tilde{\theta}_{av}(t)\right) \\ &\quad - \tfrac{k}{T}\int_0^T \cos(2\omega\tau)d\tau \cdot q'\left(\theta^* + \tilde{\theta}_{av}(t)\right) \\ &\quad + \tfrac{2k}{aT}\int_0^T \sin(\omega\tau)d\tau \cdot q\left(\theta^* + \tilde{\theta}_{av}(t)\right) \\ &= kq'\left(\theta^* + \tilde{\theta}_{av}(t)\right) \end{aligned} \tag{15}$$

Note that here $q\left(\theta^* + \tilde{\theta}_{av}(t)\right)$ and $q'\left(\theta^* + \tilde{\theta}_{av}(t)\right)$ are approximated as constants and outside the integral. The system (15) is gradient-based and locally stable under the condition (3) in Assumption 1, which is easily seen from the following fact: if $\tilde{\theta}_{av}(t) > 0$, then the gradient $q'(\cdot) < 0$ and the derivative $\dot{\tilde{\theta}}_{av}(t) < 0$, which indicates $\tilde{\theta}_{av}(t)$ decreases to zero, whereas if $\tilde{\theta}_{av}(t) < 0$, then the gradient $q'(\cdot) > 0$ and the derivative $\dot{\tilde{\theta}}_{av}(t) > 0$, which infers $\tilde{\theta}_{av}(t)$ increases to zero.

The averaged system (15) is an approximation of the original system (14). In what sense the behavior of the averaged system imitates the behavior of the original system? The problem is essential, as discussed in Khalil (2002, Chapter 10.4), but the answer is perhaps not straightforward by just comparing (15) with (14). ■

Note that $H(t) = q''\left(\theta^* + \tilde{\theta}(t) + \xi(t)a\sin(\omega t)\right)$ in (14) arises from the Taylor remainder in (12). It is close to the Hessian $H = q''(\theta^*)$ in the extremum point when both $\tilde{\theta}(t)$ and a are small. Then, the ES closed-loop error system (14) can be further rewritten as

$$\begin{aligned}\dot{\tilde{\theta}}(t) &= kq'\left(\theta^* + \tilde{\theta}(t)\right) - kq'\left(\theta^* + \tilde{\theta}(t)\right)\cos(2\omega t) \\ &\quad + \tfrac{2k}{a} q\left(\theta^* + \tilde{\theta}(t)\right)\sin(\omega t) + kaH\sin^3(\omega t) \\ &\quad + ka\left(H(t) - H\right)\sin^3(\omega t)\end{aligned} \tag{16}$$

in which the average of $kaH\sin^3(\omega t)$ is zero. By defining

$$\begin{aligned}\mathcal{F}(t) &= -\cos(2\omega t)\, q'\left(\theta^* + \tilde{\theta}(t)\right) \\ &\quad + \tfrac{2}{a}\sin(\omega t)\, q\left(\theta^* + \tilde{\theta}(t)\right) + aH\sin^3(\omega t) \\ \mathcal{R}(t) &= \left(H(t) - H\right)\sin^3(\omega t),\end{aligned} \tag{17}$$

the system (16) is compactly expressed as

$$\dot{\tilde{\theta}}(t) = kq'\left(\theta^* + \tilde{\theta}(t)\right) + k\mathcal{F}(t) + ka\mathcal{R}(t), \quad t \geq 0 \tag{18}$$

If the map (1) is of a standard quadratic form instead of a general nonlinear form, then $\mathcal{R}(t)$ is equal to 0.

Define

$$\omega = \tfrac{2\pi}{\varepsilon} \tag{19}$$

to re-scale the dither period to be ε. Via the time-delay method to averaging, we integrate $\mathcal{F}(t)$ in (17) over one dither period $[t-\varepsilon, t]$ for $t \geq \varepsilon$, we get

$$\begin{aligned}\tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t} \mathcal{F}(\tau)d\tau &= -\tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\cos\left(\tfrac{4\pi}{\varepsilon}\tau\right) q'\left(\theta^* + \tilde{\theta}(\tau)\right)d\tau \\ &\quad + \tfrac{2}{a\varepsilon}\int_{t-\varepsilon}^{t}\sin\left(\tfrac{2\pi}{\varepsilon}\tau\right) q\left(\theta^* + \tilde{\theta}(\tau)\right)d\tau\end{aligned} \tag{20}$$

where we use $\int_{t-\varepsilon}^{t}\sin^3\left(\frac{2\pi}{\varepsilon}\tau\right)d\tau = 0$. Note that here we employ the backward integral $\int_{t-\varepsilon}^{t} d\tau$ rather than the forward integral $\int_{t}^{t+\varepsilon} d\tau$ so that we are able to get a distributed time-delay system below. Besides, different from (15), $q'\left(\theta^* + \tilde{\theta}(t)\right)$ and $q\left(\theta^* + \tilde{\theta}(t)\right)$ are not handled as "freezing" constants and put outside the integral, thus we can get a more precisely transformed system than the averaged system (15).

To deal with the 1st term on the right-hand side of (20), we use

$$\begin{aligned}&-\tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\cos\left(\tfrac{4\pi}{\varepsilon}\tau\right) q'\left(\theta^* + \tilde{\theta}(\tau)\right)d\tau \\ &= \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\cos\left(\tfrac{4\pi}{\varepsilon}\tau\right)\left[q'\left(\theta^* + \tilde{\theta}(t)\right) - q'\left(\theta^* + \tilde{\theta}(\tau)\right)\right]d\tau \\ &= \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\cos\left(\tfrac{4\pi}{\varepsilon}\tau\right)\int_{\tau}^{t} q''\left(\theta^* + \tilde{\theta}(s)\right)\dot{\tilde{\theta}}(s)ds d\tau\end{aligned} \tag{21}$$

where we employ $\int_{t-\varepsilon}^{t}\cos\left(\frac{4\pi}{\varepsilon}\tau\right)d\tau \cdot q'\left(\theta^* + \tilde{\theta}(t)\right) = 0$.

To deal with the 2nd term on the right-hand side of (20), we have

$$\begin{aligned}&\tfrac{2}{a\varepsilon}\int_{t-\varepsilon}^{t}\sin\left(\tfrac{2\pi}{\varepsilon}\tau\right) q\left(\theta^* + \tilde{\theta}(\tau)\right)d\tau \\ &= -\tfrac{2}{a\varepsilon}\int_{t-\varepsilon}^{t}\sin\left(\tfrac{2\pi}{\varepsilon}\tau\right)\left[q\left(\theta^* + \tilde{\theta}(t)\right) - q\left(\theta^* + \tilde{\theta}(\tau)\right)\right]d\tau \\ &= -\tfrac{2}{a\varepsilon}\int_{t-\varepsilon}^{t}\sin\left(\tfrac{2\pi}{\varepsilon}\tau\right)\int_{\tau}^{t} q'\left(\theta^* + \tilde{\theta}(s)\right)\dot{\tilde{\theta}}(s)ds d\tau\end{aligned} \tag{22}$$

where we use $\int_{t-\varepsilon}^{t}\sin\left(\frac{2\pi}{\varepsilon}\tau\right)d\tau \cdot q\left(\theta^* + \tilde{\theta}(t)\right) = 0$.

We define

$$G(t) = \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}(\tau - t + \varepsilon)\mathcal{F}(\tau)d\tau \tag{23}$$

and employ the relation

$$\tfrac{d}{dt}\left[\tilde{\theta}(t) - kG(t)\right] = \dot{\tilde{\theta}}(t) - k\mathcal{F}(t) + \tfrac{k}{\varepsilon}\int_{t-\varepsilon}^{t}\mathcal{F}(\tau)d\tau \tag{24}$$

where $G(t)$ is borrowed from Fridman and Shaikhet (2016) to transform the original delay-free ES error system (18) into the system with the distributed time-delay (25), which is a perturbation of the averaged system (15).

Substituting (21)–(22) into (20), and further substituting (20) into (24), we present the closed-loop error system as follows:

$$\begin{aligned}\tfrac{d}{dt}\left[\tilde{\theta}(t) - kG(t)\right] &= kq'\left(\theta^* + \tilde{\theta}(t)\right) - \tfrac{2k}{a}Y_1(t) - kY_2(t) \\ &\quad + ka\mathcal{R}(t), \quad t \geq \varepsilon\end{aligned} \tag{25}$$

where

$$\begin{aligned}Y_1(t) &= \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau}^{t}\sin\left(\tfrac{2\pi}{\varepsilon}\tau\right) q'\left(\theta^* + \tilde{\theta}(s)\right)\dot{\tilde{\theta}}(s)ds d\tau \\ Y_2(t) &= -\tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau}^{t}\cos\left(\tfrac{4\pi}{\varepsilon}\tau\right) q''\left(\theta^* + \tilde{\theta}(s)\right)\dot{\tilde{\theta}}(s)ds d\tau\end{aligned} \tag{26}$$

with $\dot{\tilde{\theta}}(t)$ defined by (18).

Denoting (Yang & Fridman, 2023)

$$z(t) = \tilde{\theta}(t) - kG(t) \tag{27}$$

the closed-loop system (25) becomes

$$\begin{aligned}\dot{z}(t) &= kq'\left(\theta^* + z(t) + kG(t)\right) \\ &\quad - \tfrac{2k}{a}Y_1(t) - kY_2(t) + ka\mathcal{R}(t), \quad t \geq \varepsilon\end{aligned} \tag{28}$$

Remark 2. Comparing (25) with (15), it is observed that the time-delay system (25) has the same dominant part $\frac{d}{dt}\tilde{\theta}(t) = kq'\left(\theta^* + \tilde{\theta}(t)\right)$ as the stable averaged system (15), and has a few of additional perturbation terms $G(t), Y_1(t), Y_2(t), ka\mathcal{R}(t)$. If $\dot{\tilde{\theta}}(t)$ is of order $O(1)$, then the integral terms $G(t)$ in (23) and $Y_1(t), Y_2(t)$ in (26) are of order $O(\varepsilon)$ (see (30)), while the term $ka\mathcal{R}(t)$ is of order $O(a)$. They are close to zero when ε and a are tuned to be small enough. Namely, the time-delay plant (25) with explicit disturbances $G(t), Y_1(t), Y_2(t), ka\mathcal{R}(t)$ is a more precise model than the classical averaged system (15) in terms of behavior description of the original ES system (18). The larger ε and a, the stronger effect of $G(t), Y_1(t), Y_2(t), ka\mathcal{R}(t)$. As a result, to find the upper bounds on ε and a that preserve the practical stability of the ES system is essential, but these theoretical bounds have not been suggested in the literature.

What is more important, unlike the conversion from (14) into (15) via averaging, we do not employ any approximation or neglect anything in the transformation from (18) into (25) via the time-delay approach. Consequently, the solution of the ES system (18) is also a solution of the time-delay system (25), and we conclude the stability of the original ES system via the stability of the time-delay system.

In addition, distinct from our previous works (Zhu & Fridman, 2022; Zhu et al., 2023) where we directly analyzed the plant (25) which is a kind of neutral type differential equation, here we introduce the variable change (27) as in Yang and Fridman (2023) and analyze the ordinary differential Eq. (28) which renders the derivation to be simpler. ■

Under the bounds (6) in Assumption 2 and the overall bound (13), the upper bounds on $\mathcal{F}(t)$, $\mathcal{R}(t)$ and $\dot{\tilde{\theta}}(t)$ are obtained from Eqs. (17) and (18) such that

$$\begin{aligned}|\mathcal{F}(t)| &\leq \left|q'\left(\theta^* + \tilde{\theta}(t)\right)\right| + \tfrac{2}{a}\left|q\left(\theta^* + \tilde{\theta}(t)\right)\right| + a|H| \\ &< q_1(\sigma) + \tfrac{2}{a}q_0(\sigma) + aq_2(\sigma) \triangleq \Delta_{\mathcal{F}} \\ |\mathcal{R}(t)| &\leq \left|q''\left(\theta^* + \tilde{\theta}(t) + \xi(t)a\sin\left(\tfrac{2\pi}{\varepsilon}t\right)\right) - q''(\theta^*)\right| \\ &< L\left|\tilde{\theta}(t) + \xi(t)a\sin\left(\tfrac{2\pi}{\varepsilon}t\right)\right| < L(\sigma + a) \triangleq \Delta_{\mathcal{R}} \\ \left|\dot{\tilde{\theta}}(t)\right| &\leq k\left|q'\left(\theta^* + \tilde{\theta}(t)\right)\right| + k|\mathcal{F}(t)| + ka|\mathcal{R}(t)| \\ &< kq_1(\sigma) + k\Delta_{\mathcal{F}} + ka\Delta_{\mathcal{R}} \triangleq \Delta_{\theta}\end{aligned} \tag{29}$$

Accordingly, from (23) and (26), we have

$$
\begin{aligned}
&|G(t)| = \tfrac{1}{\varepsilon}\left|\int_{t-\varepsilon}^{t}(\tau - t + \varepsilon)\mathcal{F}(\tau)d\tau\right| \\
&\le \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}(\tau - t+\varepsilon)d\tau \cdot \sup|\mathcal{F}(\tau)| < \tfrac{\Delta_{\mathcal{F}}}{2}\varepsilon \\
&|Y_1(t)| = \tfrac{1}{\varepsilon}\left|\int_{t-\varepsilon}^{t}\int_{\tau}^{t}\sin\left(\tfrac{2\pi}{\varepsilon}\tau\right)q'\left(\theta^* + \tilde{\theta}(s)\right)\dot{\tilde{\theta}}(s)dsd\tau\right| \\
&< \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau}^{t} dsd\tau\cdot q_1(\sigma)\Delta_\theta = \tfrac{q_1(\sigma)\Delta_\theta}{2}\varepsilon \\
&|Y_2(t)| = \tfrac{1}{\varepsilon}\left|\int_{t-\varepsilon}^{t}\int_{\tau}^{t}\cos\left(\tfrac{4\pi}{\varepsilon}\tau\right)q''\left(\theta^* + \tilde{\theta}(s)\right)\dot{\tilde{\theta}}(s)dsd\tau\right| \\
&< \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau}^{t} dsd\tau\cdot q_2(\sigma)\Delta_\theta = \tfrac{q_2(\sigma)\Delta_\theta}{2}\varepsilon
\end{aligned}
\tag{30}
$$

Theorem 1. *Under Assumptions 1–2 with a given $\sigma > 0$, consider the closed-loop system consisting of the single-variable map (1) and the ES controller (11), as well as the initial condition $\left|\tilde{\theta}(0)\right| \le \sigma_0 < \sigma$. Given tuning parameters $\sigma_0, k, a, \delta, \varepsilon^* > 0$ where a and ε^* are small, let scalar decision variables $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5 > 0$ satisfy the LMIs:*

$$
\begin{aligned}
&\Omega_1 = \begin{bmatrix}
-2(k\mu(\sigma)-\delta) & 0 & -\frac{2k}{a} & -k & k & ka \\
* & -\frac{\lambda_1}{\varepsilon^*}+\lambda_5 k^2 q_2^2(\sigma) & 0 & 0 & 0 & 0 \\
* & * & -\frac{\lambda_2}{\varepsilon^*} & 0 & 0 & 0 \\
* & * & * & -\frac{\lambda_3}{\varepsilon^*} & 0 & 0 \\
* & * & * & * & -\lambda_5 & 0 \\
* & * & * & * & * & -\lambda_4 a
\end{bmatrix} \\
&< 0, \\
&\Omega_2 = \sigma_0 + (\Delta_\theta + k\Delta_{\mathcal{F}})\varepsilon^* < \sigma, \\
&\Omega_3 = \frac{\left(\lambda_1\Delta_{\mathcal{F}}^2 + \lambda_2 q_1^2(\sigma)\Delta_\theta^2 + \lambda_3 q_2^2(\sigma)\Delta_\theta^2\right)\varepsilon^* + 4\lambda_4 a\Delta_{\mathcal{R}}^2}{8\delta} \\
&< \left(\sigma - \tfrac{k\Delta_{\mathcal{F}}}{2}\varepsilon^*\right)^2
\end{aligned}
\tag{31}
$$

Then, $\forall\varepsilon \in (0, \varepsilon^*]$, the estimation error satisfies

$$
\begin{aligned}
&\left|\tilde{\theta}(t)\right| \le \left|\tilde{\theta}(0)\right| + \Delta_\theta t < \sigma, \quad t\in[0,\varepsilon], \\
&\left|\tilde{\theta}(t)\right| \le \Bigg[\left(\left|\tilde{\theta}(\varepsilon)\right| + \tfrac{k\Delta_{\mathcal{F}}}{2}\varepsilon\right)^2 e^{-2\delta(t-\varepsilon)} + \left(1 - e^{-2\delta(t-\varepsilon)}\right) \\
&\times \frac{\left(\lambda_1\Delta_{\mathcal{F}}^2 + \lambda_2 q_1^2(\sigma)\Delta_\theta^2 + \lambda_3 q_2^2(\sigma)\Delta_\theta^2\right)\varepsilon + 4\lambda_4 a\Delta_{\mathcal{R}}^2}{8\delta}\Bigg]^{\frac{1}{2}} \\
&+ \tfrac{k\Delta_{\mathcal{F}}}{2}\varepsilon < \sigma, \quad t \in [\varepsilon, \infty)
\end{aligned}
\tag{32}
$$

and is exponentially attracted to the ball

$$
\begin{aligned}
\Theta = \Bigg\{ &\tilde{\theta}\in\mathbb{R}^2 : \left|\tilde{\theta}\right| \le \tfrac{k\Delta_{\mathcal{F}}}{2}\varepsilon \\
&+ \sqrt{\frac{\left(\lambda_1\Delta_{\mathcal{F}}^2 + \lambda_2 q_1^2(\sigma)\Delta_\theta^2 + \lambda_3 q_2^2(\sigma)\Delta_\theta^2\right)\varepsilon + 4\lambda_4 a\Delta_{\mathcal{R}}^2}{8\delta}}\Bigg\}
\end{aligned}
\tag{33}
$$

which is adjustable via the tuning parameters ε and a. Moreover, LMIs (31) are always feasible for small enough ε^* and a.

The proof of Theorem 1 follows the argument of the proof of Theorem 2 given in the Appendix.

As revealed in the proof of the Appendix, in (28), the perturbation terms $Y_1(t)$, $Y_2(t)$ and $\mathcal{R}(t)$ are bounded and further compensated separately in the Lyapunov-based analysis. An alternative way is to group $Y_1(t)$, $Y_2(t)$ and $\mathcal{R}(t)$ as a whole and not to compensate it in the Lyapunov-based analysis.

Next, we present the alternatively simple calculation, but the exponential convergence about the transient performance like (32) cannot be suggested by this way.

Defining

$$
\begin{aligned}
P(t) &= q'\left(\theta^* + z(t) + kG(t)\right) - q'\left(\theta^* + z(t)\right) \\
&\quad - \tfrac{2}{a}Y_1(t) - Y_2(t) + a\mathcal{R}(t)
\end{aligned}
\tag{34}
$$

the system (28) can be presented as

$$
\dot{z}(t) = kq'\left(\theta^* + z(t)\right) + kP(t), \quad t \ge \varepsilon \tag{35}
$$

Using differential mean-value theorem, the 1st line on the right-hand side of (34) is handled as follows:

$$
\begin{aligned}
&q'\left(\theta^* + z(t) + kG(t)\right) - q'\left(\theta^* + z(t)\right) \\
&= q''\left(\theta^* + z(t) + \nu kG(t)\right)\cdot kG(t)
\end{aligned}
\tag{36}
$$

where $\nu \in (0, 1)$.

Making use of (30) and (36), the upper bound of $P(t)$ is derived as

$$
\begin{aligned}
|P(t)| &\le \left|q'\left(\theta^* + z(t) + kG(t)\right) - q'\left(\theta^* + z(t)\right)\right| \\
&\quad + \tfrac{2}{a}\left|Y_1(t)\right| + \left|Y_2(t)\right| + a\left|\mathcal{R}(t)\right| \\
&\le \sup\left|q''\left(\theta^* + z(t) + \nu kG(t)\right)\right|\cdot|kG(t)| \\
&\quad + \tfrac{2}{a}\left|Y_1(t)\right| + \left|Y_2(t)\right| + a\left|\mathcal{R}(t)\right| \\
&< \tfrac{kq_2(\sigma)\Delta_{\mathcal{F}}}{2}\varepsilon + \tfrac{q_1(\sigma)\Delta_\theta}{a}\varepsilon + \tfrac{q_2(\sigma)\Delta_\theta}{2}\varepsilon + a\Delta_{\mathcal{R}} \triangleq \Delta_P
\end{aligned}
\tag{37}
$$

Lyapunov candidate is selected as $V(t) = \frac{1}{2}z^2(t)$. Taking its derivative along (35), under (3) in Assumption 1, we have

$$
\begin{aligned}
\dot{V}(t) &= kq'\left(\theta^* + z(t)\right)\cdot z(t) + kP(t)\cdot z(t) \\
&< -k\mu(\sigma)z^2(t) + k\left|z(t)\right|\Delta_P, \quad t \ge \varepsilon
\end{aligned}
\tag{38}
$$

When $|z(t)| \ge \frac{\Delta_P}{\mu(\sigma)}$, $\dot{V}(t)$ is negative, which means that $z(t)$ decreases to the set $\left\{|z(t)| < \frac{\Delta_P}{\mu(\sigma)}\right\}$. In order to keep the overall bound $\left|\tilde{\theta}(t)\right| < \sigma$ assumed in (13), we need:

$$
\begin{cases}
\left|\tilde{\theta}(t)\right| = \left|\tilde{\theta}(0) + \int_0^t \dot{\tilde{\theta}}(\tau)d\tau\right| < \left|\tilde{\theta}(0)\right| + \Delta_\theta\varepsilon < \sigma, \\
\quad t\in[0,\varepsilon] \\
\left|\tilde{\theta}(t)\right| = |z(t) + kG(t)| < \sup_{t\ge\varepsilon}|z(t)| + \frac{k\Delta_{\mathcal{F}}}{2}\varepsilon < \sigma, \\
\quad t \in [\varepsilon, \infty)
\end{cases}
\tag{39}
$$

The above inequalities can be guaranteed by the linear inequalities for finding ε^*:

$$
\begin{cases}
\sigma_0 + (\Delta_\theta + k\Delta_{\mathcal{F}})\varepsilon^* < \sigma, \\
\dfrac{\left[\frac{kq_2(\sigma)\Delta_{\mathcal{F}}}{2} + \left(\frac{q_1(\sigma)}{a} + \frac{q_2(\sigma)}{2}\right)\Delta_\theta\right]\varepsilon^* + a\Delta_{\mathcal{R}}}{\mu(\sigma)} + \dfrac{k\Delta_{\mathcal{F}}}{2}\varepsilon^* < \sigma
\end{cases}
\tag{40}
$$

and the attractive ball for the estimation error has the form:

$$
\begin{aligned}
\Theta = \Bigg\{ &\tilde{\theta}\in\mathbb{R} : \left|\tilde{\theta}\right| < \Bigg\{\left[\tfrac{kq_2(\sigma)\Delta_{\mathcal{F}}}{2} + \left(\tfrac{q_1(\sigma)}{a} + \tfrac{q_2(\sigma)}{2}\right)\Delta_\theta\right]\varepsilon \\
&+ a\Delta_{\mathcal{R}}\Bigg\}\cdot\tfrac{1}{\mu(\sigma)} + \tfrac{k\Delta_{\mathcal{F}}}{2}\varepsilon\Bigg\}, \quad \varepsilon \in (0, \varepsilon^*].
\end{aligned}
\tag{41}
$$

Proposition 1. *Under Assumptions 1–2 with a given $\sigma > 0$, consider the closed-loop system consisting of the single-variable map (1) and the ES controller (11), as well as the initial condition $\left|\tilde{\theta}(0)\right| \le \sigma_0 < \sigma$. Given tuning parameters $k, a, \varepsilon^* > 0$ and σ_0, let the linear inequalities (40) with respect to ε^* are satisfied. Then, $\forall\varepsilon \in (0, \varepsilon^*]$, the estimation error satisfies (39), and is attracted to the ball (41).*

To illustrate that (39) is implied by (40), we consider two different situations. Firstly, if the initial condition $|z(\varepsilon)| < \frac{\Delta_P}{\mu(\sigma)}$, then $z(t)$ for $t \ge \varepsilon$ stays in $\left\{|z(t)| < \frac{\Delta_P}{\mu(\sigma)}\right\}$ and from the 2nd equation of (39) we have $\left|\tilde{\theta}(t)\right| < \frac{\Delta_P}{\mu(\sigma)} + \frac{k\Delta_{\mathcal{F}}}{2}\varepsilon < \sigma$ which holds due to the 2nd inequality of (40). If $\frac{\Delta_P}{\mu(\sigma)} \le |z(\varepsilon)| < \sigma - \frac{k\Delta_{\mathcal{F}}}{2}\varepsilon$, then $|z(t)|$ for $t \ge \varepsilon$ monotonically decreases due to (38) until it reaches the region $|z(t)| < \frac{\Delta_P}{\mu(\sigma)}$. Then when $t \ge \varepsilon$, we have $\left|\tilde{\theta}(t)\right| \le |z(t)| + k\left|G(t)\right| < |z(\varepsilon)| + \frac{k\Delta_{\mathcal{F}}}{2}\varepsilon < \left|\tilde{\theta}(\varepsilon)\right| + k\Delta_{\mathcal{F}}\varepsilon < \sigma$ which is ensured by the 1st inequality of (40).

Remark 3. Comparing (14) with (16), we utilize $kaH(t)\sin^3(\omega t) = kaH\sin^3(\omega t) + ka\left(H(t) - H\right)\sin^3(\omega t)$, in which the average of the former is zero, and $|H(t)-H|$ in the latter is upper bounded by (29) under Lipschitz condition.

An alternative bound to (29) without Lipschitz assumption is that $|H(t)| < |q_2(\sigma)|$ under (6), which is more conservative in terms of LMI-based analysis. By an empirical study via simulation examples in Section 4, we have $\sup_{t\geq 0}|H(t)-H| \ll \sup_{t\geq 0}|H(t)|$. ■

3. Vector systems

In this section, we extend the time-delay approach from scalar plants to vector plants, which is non-trivial. For notational simplicity, we address the case of two variables. The method could be generalized to cases with any n variables in a recursive way, but the calculation is much longer.

Consider the two-variable static maps given by

$$y(t) = q\left(\theta(t)\right) \tag{42}$$

where $y(t) \in \mathbb{R}$ is the measurable output and $\theta(t) = [\theta_1(t), \theta_2(t)]^T \in \mathbb{R}^2$ is the control input vector. The map (42) meets the following assumptions.

Assumption 3. There exist a constant vector $\theta^* = \left[\theta_1^*, \theta_2^*\right]^T \in \mathbb{R}^2$, constants $\sigma > 0$ and $a > 0$, the map $q(\theta) = q(\theta_1, \theta_2) \in C^2\left[\left(\theta_1^*-\sigma-a, \theta_1^*+\sigma+a\right) \times \left(\theta_2^*-\sigma-a, \theta_2^*+\sigma+a\right)\right]$, and the following hold:

$$\begin{aligned}
&\frac{\partial q}{\partial \theta}(\theta^*) = \left[\begin{array}{cc}\frac{\partial q}{\partial \theta_1}(\theta^*) & \frac{\partial q}{\partial \theta_2}(\theta^*)\end{array}\right] = 0,\\
&\frac{\partial^2 q}{\partial \theta^2}(\theta^*) = \begin{bmatrix}\frac{\partial^2 q}{\partial \theta_1^2}(\theta^*) & \frac{\partial^2 q}{\partial \theta_1 \partial \theta_2}(\theta^*)\\ * & \frac{\partial^2 q}{\partial \theta_2^2}(\theta^*)\end{bmatrix} = H < 0,
\end{aligned} \tag{43}$$

$$\begin{aligned}
&\frac{\partial q}{\partial \theta}(\theta^*+\Delta)\cdot\Delta \leq -\mu(\sigma)\cdot|\Delta|^2 < 0,\\
&\Delta = [\Delta_1, \Delta_2]^T, \quad \forall 0 < |\Delta_i| < \sigma, \quad i = 1, 2
\end{aligned} \tag{44}$$

where $\mu(\sigma) > 0$ is a σ-dependent constant.

Assumption 4. For any $\Delta = [\Delta_1, \Delta_2]^T$ with $0 < |\Delta_i| < \sigma$, $i = 1, 2$ and a defined in Assumption 3, given $\xi = [\xi_1, \xi_2]^T$ with $\xi_1, \xi_2 \in [-1, 1]$, we have

$$\begin{aligned}
&|q(\theta^*+\Delta)| < q_0(\sigma),\\
&\left|\frac{\partial q}{\partial \theta}(\theta^*+\Delta)\right| < q_1(\sigma), \quad \left|\frac{\partial^2 q}{\partial \theta^2}(\theta^*+\Delta)\right| < q_2(\sigma),\\
&\left|\frac{\partial^2 q}{\partial \theta^2}(\theta^*+\Delta+a\xi) - \frac{\partial^2 q}{\partial \theta^2}(\theta^*)\right| < L|\Delta+a\xi|
\end{aligned} \tag{45}$$

where $q_0(\sigma)$, $q_1(\sigma)$, $q_2(\sigma)$, L are known positive constants.

The condition (44) is not restrictive as shown in the following calculation.

$$\begin{aligned}
&\frac{\partial q}{\partial \theta}(\theta^*+\Delta)\cdot\Delta\\
&= \left[\begin{array}{cc}\frac{\partial q}{\partial \theta_1}(\theta_1^*+\Delta_1, \theta_2^*+\Delta_2) & \frac{\partial q}{\partial \theta_2}(\theta_1^*+\Delta_1, \theta_2^*+\Delta_2)\end{array}\right]\cdot\left[\begin{array}{c}\Delta_1\\ \Delta_2\end{array}\right]\\
&= \frac{\partial q}{\partial \theta_1}\left(\theta_1^*+\Delta_1, \theta_2^*+\Delta_2\right)\cdot\Delta_1\\
&\quad + \frac{\partial q}{\partial \theta_2}\left(\theta_1^*+\Delta_1, \theta_2^*+\Delta_2\right)\cdot\Delta_2
\end{aligned} \tag{46}$$

Define two functions $f_1(\xi)$, $f_2(\xi)$, $\xi \in [0, 1]$ as

$$\begin{cases}
f_1(\xi) = \dfrac{\partial q}{\partial \theta_1}\left(\theta_1^*+\xi\Delta_1, \theta_2^*+\xi\Delta_2\right)\\
f_2(\xi) = \dfrac{\partial q}{\partial \theta_2}\left(\theta_1^*+\xi\Delta_1, \theta_2^*+\xi\Delta_2\right)
\end{cases} \tag{47}$$

Under the fact that

$$\begin{aligned}
&f_1(0) = f_2(0) = 0\\
&f_1(1) = f_1(0) + \int_0^1 f_1'(\xi)d\xi = \int_0^1 f_1'(\xi)d\xi\\
&f_2(1) = f_2(0) + \int_0^1 f_2'(\xi)d\xi = \int_0^1 f_2'(\xi)d\xi
\end{aligned} \tag{48}$$

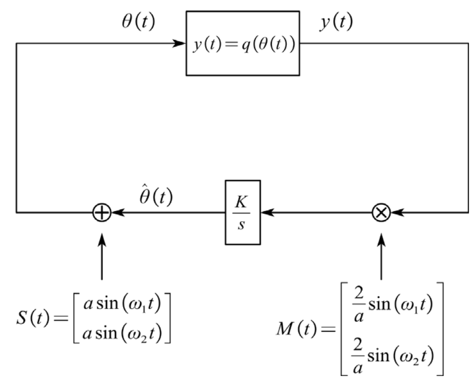


Fig. 3. ES for a two-variable static map.

the formula (46) is expressed as

$$\begin{aligned}
&\frac{\partial q}{\partial \theta}(\theta^*+\Delta)\cdot\Delta = f_1(1)\Delta_1 + f_2(1)\Delta_2\\
&= \int_0^1 \left[\Delta_1\cdot f_1'(\xi) + \Delta_2\cdot f_2'(\xi)\right]d\xi
\end{aligned} \tag{49}$$

Furthermore,

$$\begin{aligned}
f_1'(\xi) &= \frac{\partial^2 q}{\partial \theta_1^2}\left(\theta_1^*+\xi\Delta_1, \theta_2^*+\xi\Delta_2\right)\cdot\Delta_1\\
&\quad + \frac{\partial^2 q}{\partial \theta_1 \partial \theta_2}\left(\theta_1^*+\xi\Delta_1, \theta_2^*+\xi\Delta_2\right)\cdot\Delta_2\\
f_2'(\xi) &= \frac{\partial^2 q}{\partial \theta_2 \partial \theta_1}\left(\theta_1^*+\xi\Delta_1, \theta_2^*+\xi\Delta_2\right)\cdot\Delta_1\\
&\quad + \frac{\partial^2 q}{\partial \theta_2^2}\left(\theta_1^*+\xi\Delta_1, \theta_2^*+\xi\Delta_2\right)\cdot\Delta_2
\end{aligned} \tag{50}$$

Substituting (50) into (49) and using $(\theta^*+\xi\Delta) = \left(\theta_1^*+\xi\Delta_1, \theta_2^*+\xi\Delta_2\right)$ for notational simplicity, we get

$$\begin{aligned}
&\frac{\partial q}{\partial \theta}(\theta^*+\Delta)\cdot\Delta = \int_0^1 \left[\frac{\partial^2 q}{\partial \theta_1^2}(\theta^*+\xi\Delta)\cdot\Delta_1^2\right.\\
&\left.\quad + 2\frac{\partial^2 q}{\partial \theta_1 \partial \theta_2}(\theta^*+\xi\Delta)\cdot\Delta_1\Delta_2 + \frac{\partial^2 q}{\partial \theta_2^2}(\theta^*+\xi\Delta)\cdot\Delta_2^2\right]d\xi\\
&= \int_0^1 \left[\begin{array}{cc}\Delta_1 & \Delta_2\end{array}\right]\cdot\begin{bmatrix}\frac{\partial^2 q}{\partial \theta_1^2}(\theta^*+\xi\Delta) & \frac{\partial^2 q}{\partial \theta_1 \partial \theta_2}(\theta^*+\xi\Delta)\\ * & \frac{\partial^2 q}{\partial \theta_2^2}(\theta^*+\xi\Delta)\end{bmatrix}\cdot\left[\begin{array}{c}\Delta_1\\ \Delta_2\end{array}\right]d\xi
\end{aligned} \tag{51}$$

The inequality $\frac{\partial^2 q}{\partial \theta^2}(\theta^*) < 0$ and the continuity of $\frac{\partial^2 q}{\partial \theta^2}(\theta)$ in the domain $\left[\left(\theta_1^*-\sigma, \theta_1^*+\sigma\right) \times \left(\theta_2^*-\sigma, \theta_2^*+\sigma\right)\right]$ indicates that there always exists the deviation Δ to let the average from $\frac{\partial^2 q}{\partial \theta^2}(\theta^*)$ to $\frac{\partial^2 q}{\partial \theta^2}(\theta^*+\Delta)$ is negative definite. Thus,

$$\begin{aligned}
&\int_0^1 \left[\begin{array}{cc}\Delta_1 & \Delta_2\end{array}\right]\cdot\begin{bmatrix}\frac{\partial^2 q}{\partial \theta_1^2}(\theta^*+\xi\Delta) & \frac{\partial^2 q}{\partial \theta_1 \partial \theta_2}(\theta^*+\xi\Delta)\\ * & \frac{\partial^2 q}{\partial \theta_2^2}(\theta^*+\xi\Delta)\end{bmatrix}\cdot\left[\begin{array}{c}\Delta_1\\ \Delta_2\end{array}\right]d\xi\\
&\leq \left[\begin{array}{cc}\Delta_1 & \Delta_2\end{array}\right]\left[-\mu(\sigma)\cdot I\right]\left[\begin{array}{c}\Delta_1\\ \Delta_2\end{array}\right] = -\mu(\sigma)\cdot|\Delta|^2
\end{aligned} \tag{52}$$

Then, (44) holds.

To make $\theta(t) \to \theta^*$, the on-line estimate $\hat{\theta}(t)$ is introduced and the estimation error $\tilde{\theta}(t)$ is defined below,

$$\tilde{\theta}(t) = \hat{\theta}(t) - \theta^* \tag{53}$$

The gradient-based ES algorithm shown in Fig. 3 is employed as follows:

$$\begin{aligned}
&\theta(t) = \hat{\theta}(t) + S(t)\\
&\dot{\hat{\theta}}(t) = K\cdot M(t)\cdot y(t)
\end{aligned} \tag{54}$$

where $K = \left[\begin{smallmatrix}k & \\ & k\end{smallmatrix}\right]$ is a diagonal matrix of the adaptation gain whose sign is opposite to the sign of the Hessian matrix H in (43), $S(t) = a\left[\begin{smallmatrix}\sin(\omega_1 t)\\ \sin(\omega_2 t)\end{smallmatrix}\right]$ is the additive perturbation, $M(t) = \frac{2}{a}\left[\begin{smallmatrix}\sin(\omega_1 t)\\ \sin(\omega_2 t)\end{smallmatrix}\right]$ is the multiplicative demodulation. The condition $\omega_1 \neq \omega_2$ and $\frac{\omega_1}{\omega_2}$

is rational is requested. The initial value satisfies $\left|\hat{\theta}(0)-\theta^*\right| < \sigma_0$, where $\sigma_0 < \sigma$ is a known constant.

We consider Taylor expansion of the map (42) such that

$$\begin{aligned} &q(\theta(t)) = q\left(\theta^* + \tilde{\theta}(t) + S(t)\right) \\ &= q\left(\theta^* + \tilde{\theta}(t)\right) + \tfrac{\partial q}{\partial \theta}\left(\theta^* + \tilde{\theta}(t)\right) S(t) \\ &\quad + \tfrac{1}{2} S^T(t) H(t) S(t) \end{aligned} \tag{55}$$

where $H(t) = \frac{\partial^2 q}{\partial \theta^2}\left(\theta^* + \tilde{\theta}(t) + \xi(t) S(t)\right)$ and $\xi(t) \in (0, 1)$.

Substituting (55) into the 2nd equation of (54), and utilizing the notation $S_0(t) = \begin{bmatrix} \sin(\omega_1 t) \\ \sin(\omega_2 t) \end{bmatrix}$, $S(t) = aS_0(t)$, $M(t) = \frac{2}{a} S_0(t)$, we arrive at the dynamics of the estimation error of (53) such that

$$\begin{aligned} \dot{\tilde{\theta}}(t) &= \tfrac{2k}{a} S_0(t) q\left(\theta^* + \tilde{\theta}(t)\right) + kaS_0(t) S_0^T(t) H(t) S_0(t) \\ &\quad + 2kS_0(t) \tfrac{\partial q}{\partial \theta}\left(\theta^* + \tilde{\theta}(t)\right) S_0(t) \end{aligned} \tag{56}$$

For the last term on the right-hand side of (56), we have $S_0(t)\frac{\partial q}{\partial \theta}\left(\theta^* + \tilde{\theta}(t)\right) S_0(t) = S_0(t) S_0^T(t) \frac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(t)\right)$, and

$$\begin{aligned} &2S_0(t) S_0^T(t) = \begin{bmatrix} 2\sin^2(\omega_1 t) & 2\sin(\omega_1 t)\sin(\omega_2 t) \\ 2\sin(\omega_1 t)\sin(\omega_2 t) & 2\sin^2(\omega_2 t) \end{bmatrix} \\ &= I + L_0 \end{aligned} \tag{57}$$

where I is the unit matrix and

$$L_0(t) = \begin{bmatrix} -\cos(2\omega_1 t) & 2\sin(\omega_1 t)\sin(\omega_2 t) \\ 2\sin(\omega_1 t)\sin(\omega_2 t) & -\cos(2\omega_2 t) \end{bmatrix} \tag{58}$$

Thus, (56) becomes

$$\begin{aligned} \dot{\tilde{\theta}}(t) &= \tfrac{2k}{a} S_0(t) q\left(\theta^* + \tilde{\theta}(t)\right) + kaS_0(t) S_0^T(t) H(t) S_0(t) \\ &\quad + k \tfrac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(t)\right) + kL_0(t) \tfrac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(t)\right) \end{aligned} \tag{59}$$

Note that the Taylor remainder from (55) $H(t) = \frac{\partial^2 q}{\partial \theta^2}\left(\theta^* + \tilde{\theta}(t) + a\xi(t) S_0(t)\right)$ is close to $H = \frac{\partial^2 q}{\partial \theta^2}(\theta^*)$ if both $\tilde{\theta}(t)$ and a are small, so that $\sup_{t\geq 0}|H(t) - H| \ll \sup_{t\geq 0}|H(t)|$. Hence, we define

$$\begin{aligned} \mathcal{F}(t) &= \tfrac{2}{a} S_0(t) q\left(\theta^* + \tilde{\theta}(t)\right) + L_0(t) \tfrac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(t)\right) \\ &\quad + aS_0(t) S_0^T(t) H S_0(t) \\ \mathcal{R}(t) &= S_0(t) S_0^T(t) \left(H(t) - H\right) S_0(t) \end{aligned} \tag{60}$$

and (59) takes the form of

$$\dot{\tilde{\theta}}(t) = k \tfrac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(t)\right) + k\mathcal{F}(t) + ka\mathcal{R}(t), \quad t \geq 0 \tag{61}$$

Set

$$\omega_1 = \tfrac{2\pi l_1}{\varepsilon}, \quad \omega_2 = \tfrac{2\pi l_2}{\varepsilon} \tag{62}$$

where $l_1 \neq l_2 > 0$ are integers.

Apply the time-delay approach to $\mathcal{F}(t)$, we have

$$\begin{aligned} \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t} \mathcal{F}(\tau) d\tau &= \tfrac{2}{a\varepsilon}\int_{t-\varepsilon}^{t} S_0(\tau) q\left(\theta^* + \tilde{\theta}(\tau)\right) d\tau \\ &\quad + \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t} L_0(\tau) \tfrac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(\tau)\right) d\tau \end{aligned} \tag{63}$$

where we use $\frac{a}{\varepsilon}\int_{t-\varepsilon}^{t} S_0(\tau) S_0^T(\tau) H S_0(\tau) d\tau = 0$.

Considering the 1st term on the right-hand side of (63), we get

$$\begin{aligned} &\tfrac{2}{a\varepsilon}\int_{t-\varepsilon}^{t} S_0(\tau) q\left(\theta^* + \tilde{\theta}(\tau)\right) d\tau \\ &= -\tfrac{2}{a\varepsilon}\int_{t-\varepsilon}^{t} S_0(\tau)\left[q\left(\theta^* + \tilde{\theta}(t)\right) - q\left(\theta^* + \tilde{\theta}(\tau)\right)\right] d\tau \\ &= -\tfrac{2}{a\varepsilon}\int_{t-\varepsilon}^{t} S_0(\tau)\int_{\tau}^{t} \tfrac{\partial q}{\partial \theta}\left(\theta^* + \tilde{\theta}(s)\right) \dot{\tilde{\theta}}(s) ds d\tau \end{aligned} \tag{64}$$

where we employ $\int_{t-\varepsilon}^{t} S_0(\tau) q\left(\theta^* + \tilde{\theta}(t)\right) d\tau = 0$.

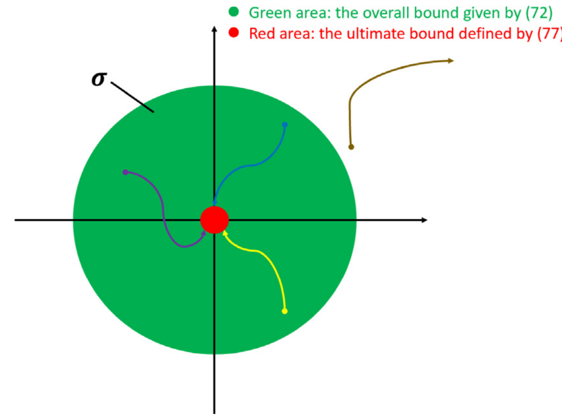


Fig. 4. Domains defined by (72) and (77). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Considering the 2nd term on the right-hand side of (63), we have

$$\begin{aligned} &\tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t} L_0(\tau) \tfrac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(\tau)\right) d\tau \\ &= -\tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t} L_0(\tau)\left[\tfrac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(t)\right) - \tfrac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(\tau)\right)\right] d\tau \\ &= -\tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t} L_0(\tau)\int_{\tau}^{t} \tfrac{\partial^2 q}{\partial \theta^2}\left(\theta^* + \tilde{\theta}(s)\right) \dot{\tilde{\theta}}(s) ds d\tau \end{aligned} \tag{65}$$

where we employ $\int_{t-\varepsilon}^{t} L_0(\tau) \frac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(t)\right) d\tau = 0$.

Now, we define

$$G(t) = \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t} (\tau - t + \varepsilon) \mathcal{F}(\tau) d\tau \tag{66}$$

and employ the relation

$$\tfrac{d}{dt}\left[\tilde{\theta}(t) - kG(t)\right] = \dot{\tilde{\theta}}(t) - k\mathcal{F}(t) + \tfrac{k}{\varepsilon}\int_{t-\varepsilon}^{t} \mathcal{F}(\tau) d\tau \tag{67}$$

and substitute (64)–(65) into (63), the closed-loop system is derived as

$$\begin{aligned} &\tfrac{d}{dt}\left[\tilde{\theta}(t) - kG(t)\right] = k \tfrac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(t)\right) \\ &\quad - \tfrac{2k}{a} Y_1(t) - kY_2(t) + ka\mathcal{R}(t), \quad t \geq \varepsilon \end{aligned} \tag{68}$$

where

$$\begin{aligned} Y_1(t) &= \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau}^{t} S_0(\tau) \tfrac{\partial q}{\partial \theta}\left(\theta^* + \tilde{\theta}(s)\right) \dot{\tilde{\theta}}(s) ds d\tau \\ Y_2(t) &= \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau}^{t} L_0(\tau) \tfrac{\partial^2 q}{\partial \theta^2}\left(\theta^* + \tilde{\theta}(s)\right) \dot{\tilde{\theta}}(s) ds d\tau \end{aligned} \tag{69}$$

Denoting

$$z(t) = \tilde{\theta}(t) - kG(t) \tag{70}$$

the system (68) becomes

$$\begin{aligned} \dot{z}(t) &= k \tfrac{\partial q}{\partial \theta}^T\left(\theta^* + z(t) + kG(t)\right) \\ &\quad - \tfrac{2k}{a} Y_1(t) - kY_2(t) + ka\mathcal{R}(t), \quad t \geq \varepsilon \end{aligned} \tag{71}$$

In parallel with (13) of the scalar case, under σ defined in Assumption 3, the overall bound for the estimation error in the vector case is supposed to satisfy

$$\left|\tilde{\theta}(t)\right| < \sigma, \quad \forall t \geq 0 \tag{72}$$

The LMI conditions (75) in Theorem 2 ensure (72).

As revealed in Fig. 4, the overall bound given by (72) (the green area in the picture) is distinct from the ultimate bound of the estimation error defined by (77) (the red area in the picture). The overall bound given by (72) represents the range in which ES can be conducted. It means that every trajectory of ES closed-loop system starting in the initial domain $\left|\tilde{\theta}(0)\right| \leq \sigma_0 < \sigma$

remains in the domain and approaches the extremum as time grows (see the curves in blue, yellow and purple inside the green area), whereas trajectories starting outside the initial domain may diverge (see the brown one outside the green area). This is consistent with the concepts of region of attraction and local stability of nonlinear systems (see Khalil (2002, Chapter 4.1)). The ultimate bound of the estimation error given by (77), describes how close the real-time estimate converges to the true extremum as $t \to \infty$. It is seen from (77) the error ultimate bound depends upon the designing parameters ε and a, and could be tuned to be sufficiently small when ε and a are chosen to be small. This result is in line with the literature (Ariyur & Krstic, 2003). The area of (77) is embraced in the area of (72). In Fig. 4, the extremum to be sought is assumed to be the origin without loss of generality, and the ES controller is brought in to render the initial estimate which is located in the green area converging to the target red area. The latter is typically much smaller.

Furthermore, by defining $\bar{S}_0 = \sup_{t\geq 0} |S_0(t)|$, $\bar{L}_0 = \sup_{t\geq 0} |L_0(t)|$, under (45), we get the upper bounds for (60) and (61) such that

$$
\begin{aligned}
&|\mathcal{F}(t)| \leq \tfrac{2}{a}\left|S_0(t)q\left(\theta^* + \tilde{\theta}(t)\right)\right| \\
&\quad + \left|L_0(t)\tfrac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(t)\right)\right| + a\left|S_0(t)S_0^T(t)HS_0(t)\right| \\
&< \tfrac{2}{a}\bar{S}_0 q_0(\sigma) + \bar{L}_0 q_1(\sigma) + a\bar{S}_0^3 q_2(\sigma) \triangleq \Delta_{\mathcal{F}} \\
&|\mathcal{R}(t)| \leq \left|S_0(t)S_0^T(t)\left(H(t) - H\right)S_0(t)\right| \\
&< \bar{S}_0^3 L\left|\tilde{\theta}(t) + a\xi(t)S_0(t)\right| < \bar{S}_0^3 L\left(\sigma + a\bar{S}_0\right) \triangleq \Delta_{\mathcal{R}} \\
&\left|\dot{\tilde{\theta}}(t)\right| \leq k\left|\tfrac{\partial q}{\partial \theta}^T\left(\theta^* + \tilde{\theta}(t)\right)\right| + k|\mathcal{F}(t)| + ka|\mathcal{R}(t)| \\
&< kq_1(\sigma) + k\Delta_{\mathcal{F}} + ka\Delta_{\mathcal{R}} \triangleq \Delta_\theta
\end{aligned}
\tag{73}
$$

Correspondingly, from (66) and (69), we have

$$
\begin{aligned}
&|G(t)| = \tfrac{1}{\varepsilon}\left|\int_{t-\varepsilon}^{t}(\tau - t + \varepsilon)\mathcal{F}(\tau)d\tau\right| \\
&\leq \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}(\tau - t + \varepsilon)d\tau \cdot \sup|\mathcal{F}(\tau)| < \tfrac{\Delta_{\mathcal{F}}}{2}\varepsilon \\
&|Y_1(t)| = \tfrac{1}{\varepsilon}\left|\int_{t-\varepsilon}^{t}\int_{\tau}^{t} S_0(\tau)\tfrac{\partial q}{\partial \theta}\left(\theta^* + \tilde{\theta}(s)\right)\dot{\tilde{\theta}}(s)dsd\tau\right| \\
&< \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau}^{t} dsd\tau \cdot \bar{S}_0 q_1(\sigma)\Delta_\theta = \tfrac{\bar{S}_0 q_1(\sigma)\Delta_\theta}{2}\varepsilon \\
&|Y_2(t)| = \tfrac{1}{\varepsilon}\left|\int_{t-\varepsilon}^{t}\int_{\tau}^{t} L_0(\tau)\tfrac{\partial^2 q}{\partial \theta^2}\left(\theta^* + \tilde{\theta}(s)\right)\dot{\tilde{\theta}}(s)dsd\tau\right| \\
&< \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau}^{t} dsd\tau \cdot \bar{L}_0 q_2(\sigma)\Delta_\theta = \tfrac{\bar{L}_0 q_2(\sigma)\Delta_\theta}{2}\varepsilon
\end{aligned}
\tag{74}
$$

Theorem 2. *Under Assumptions 3–4 with a given $\sigma > 0$, consider the closed-loop system consisting of the multi-variable map (42) and the ES controller (54), as well as the initial condition $\left|\tilde{\theta}(0)\right| \leq \sigma_0 < \sigma$. Given tuning parameters $\sigma_0, k, a, \delta, \varepsilon^* > 0$ where a and ε^* are small, let decision variables $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5 > 0$ satisfy the LMIs:*

$$
\begin{aligned}
&\Omega_1 = \\
&\begin{bmatrix}
-2(k\mu(\sigma)-\delta)I & 0 & -\frac{2k}{a}I & -kI & kI & kaI \\
* & -\left(\frac{\lambda_1}{\varepsilon^*} - \lambda_5 k^2 q_2^2(\sigma)\right)I & 0 & 0 & 0 & 0 \\
* & * & -\frac{\lambda_2}{\varepsilon^*}I & 0 & 0 & 0 \\
* & * & * & -\frac{\lambda_3}{\varepsilon^*}I & 0 & 0 \\
* & * & * & * & -\lambda_5 I & 0 \\
* & * & * & * & * & -\lambda_4 aI
\end{bmatrix} \\
&< 0, \\
&\Omega_2 = \sigma_0 + \left(\Delta_\theta + k\Delta_{\mathcal{F}}\right)\varepsilon^* < \sigma, \\
&\Omega_3 = \frac{\left(\lambda_1\Delta_{\mathcal{F}}^2 + \left(\lambda_2\bar{S}_0^2 q_1^2(\sigma) + \lambda_3\bar{L}_0^2 q_2^2(\sigma)\right)\Delta_\theta^2\right)\varepsilon^* + 4\lambda_4 a\Delta_{\mathcal{R}}^2}{8\delta} \\
&< \left(\sigma - \tfrac{k\Delta_{\mathcal{F}}}{2}\varepsilon^*\right)^2
\end{aligned}
\tag{75}
$$

Then, $\forall \varepsilon \in (0, \varepsilon^*]$, the estimation error satisfies

$$
\begin{aligned}
&\left|\tilde{\theta}(t)\right| \leq \left|\tilde{\theta}(0)\right| + \Delta_\theta t < \sigma, \quad t \in [0, \varepsilon] \\
&\left|\tilde{\theta}(t)\right| \leq \Bigg[\left(\left|\tilde{\theta}(\varepsilon)\right| + \tfrac{k\Delta_{\mathcal{F}}}{2}\varepsilon\right)^2 e^{-2\delta(t-\varepsilon)} + \left(1 - e^{-2\delta(t-\varepsilon)}\right) \\
&\quad \times \frac{\left(\lambda_1\Delta_{\mathcal{F}}^2 + \lambda_2\bar{S}_0^2 q_1^2(\sigma)\Delta_\theta^2 + \lambda_3\bar{L}_0^2 q_2^2(\sigma)\Delta_\theta^2\right)\varepsilon + 4\lambda_4 a\Delta_{\mathcal{R}}^2}{8\delta}\Bigg]^{\frac{1}{2}} \\
&\quad + \tfrac{k\Delta_{\mathcal{F}}}{2}\varepsilon < \sigma, \quad t \in [\varepsilon, \infty)
\end{aligned}
\tag{76}
$$

and is exponentially attracted to the ball

$$
\begin{aligned}
\Theta = \Bigg\{ &\tilde{\theta} \in \mathbb{R}^2 : \left|\tilde{\theta}\right| \leq \tfrac{k\Delta_{\mathcal{F}}}{2}\varepsilon \\
&+ \sqrt{\frac{\left(\lambda_1\Delta_{\mathcal{F}}^2 + \lambda_2\bar{S}_0^2 q_1^2(\sigma)\Delta_\theta^2 + \lambda_3\bar{L}_0^2 q_2^2(\sigma)\Delta_\theta^2\right)\varepsilon + 4\lambda_4 a\Delta_{\mathcal{R}}^2}{8\delta}} \Bigg\}
\end{aligned}
\tag{77}
$$

which is adjustable by the tuning parameters ε and a. Moreover, LMIs (75) are always feasible for small enough ε^* and a.

See the Appendix for proof.

Table 1
Real: $q_0 = 3.8610$, $q_1 = 6.2900$, $q_2 = 5.4000$, $L = 2.0$.

	ε	δ	σ_0	σ	$\mu(\sigma)$	UB
Th1	0.0001	0.000248	1.6	1.7	0.3	0.4287
Prop1	0.0001	–	1.6	1.7	0.3	0.2719

Table 2
Estimated: $q_0 = 5.2297$, $q_1 = 7.4100$, $q_2 = 5.8000$, $L = 2.05$.

	ε	δ	σ_0	σ	$\mu(\sigma)$	UB
Th1	0.0001	0.000264	1.6	1.7	0.3	0.5897
Prop1	0.0001	–	1.6	1.7	0.3	0.3026

4. Examples

4.1. Scalar case

Given the nonlinear scalar map (7)

$$
q(\theta) = \tfrac{1}{3}\theta^3 - \theta \tag{78}
$$

with the ES controller (11) with $k = 0.001$, $a = 0.02$. The upper bounds in (6) are available. We perform the simulations with both the LMIs (31) in Theorem 1 and the inequalities (40) in Proposition 1. We provide two groups of data: For a given σ, the "real" bounds refer to the exact bounds calculated by (6) as if we know the map exactly, whereas the "estimated" bounds refer to the approximate bounds which are somewhat larger than the real bounds. The solutions for both real bounds and estimated bounds via Theorem 1 and Proposition 1 are shown in Tables 1–2, respectively, where "UB" refers to the ultimate bound $\lim_{t\to\infty}\sup\left|\tilde{\theta}(t)\right|$. Note that the tuning parameters specified by designers are the controller gain k, the dither period ε and the dither amplitude a. In (31), λ_1, λ_2, λ_3, λ_4 and λ_5 are just decision variables of LMIs, whose calculations can be left up to computers, if the designing parameters are given. We have tuned k and a to maximize ε.

It is seen from Table 1 that if $\sup_{t\geq 0}|H(t) - H| = L(\sigma + a) = 3.44$ in (29) is replaced by $\sup_{t\geq 0}|H(t)| = q_2 = 5.4$, it results in $\varepsilon = 0.000015$ and $a = 0.01$ which is more conservative. The same conclusion holds for Table 2, verifying the content in Remark 3.

Table 3
Real: $q_0 = 1.1668$, $q_1 = 1.8857$, $q_2 = 2$, $L = 0.7778$.

	ε	δ	σ_0	σ	$\mu(\sigma)$	UB
Th2	0.01	0.0055	0.9	1.0	1.0	0.4182
Th2	0.03	0.0032	0.9	1.0	1.0	0.6072

Table 4
Estimated: $q_0 = 1.3$, $q_1 = 2$, $q_2 = 2.1$, $L = 1$.

	ε	δ	σ_0	σ	$\mu(\sigma)$	UB
Th2	0.01	0.0052	0.9	1.0	1.0	0.5165
Th2	0.03	0.0042	0.9	1.0	1.0	0.7138

4.2. Vector case

We consider a nonlinear function with two variables

$$q(x,y) = -\tfrac{1}{2}(x^2+y^2) - \tfrac{1}{12}x^2y^2 - \tfrac{1}{24}(x^4+y^4) \tag{79}$$

which has a maximum at $(0,0)$ as shown in Fig. 5. We employ the ES controller (54) with parameters: $k = 0.008$, $a = 0.1$, and $\omega_1 = \omega$, $\omega_2 = 2\omega$. The LMI solutions under both the real and estimated bounds of q_0, q_1, q_2, L in (45) for a given σ are shown in Tables 3–4, respectively. The LMI solution results in all the tables are obtained by optimizing with the "fmincon" function in Matlab. Namely, the decision variables λ_1, λ_2, λ_3, λ_4, λ_5 in Theorems 1 and 2 are optimized by the "fmincon" function to achieve a smaller ultimate bound for the estimation error. Similar to Remark 2 in Yang and Fridman (2023), we can further improve the results by setting the ultimate bound that we got each time as the new σ, and perform the simulation in an iterative way so that a much smaller ultimate bound is attainable. For the simulation under the initial condition $x(0) = 1, y(0) = -1$, and $\varepsilon = 0.03$, the ES trajectory is shown in Fig. 6.

In the existing literature (Yang & Fridman, 2023; Zhu & Fridman, 2022; Zhu et al., 2023), the constant disturbances were allowed to be added to the map of the quadratic form. However, here the 4th-power nonlinear term $-\frac{1}{12}x^2y^2 - \frac{1}{24}(x^4+y^4)$ in (79) cannot be handled as a bounded disturbance as it is state-dependent and time-varying. To be specific, the example (79) is rewritten as

$$q(x(t),y(t)) = -\tfrac{1}{2}(x^2(t)+y^2(t)) + \Delta(t) \tag{80}$$

where

$$\begin{aligned}\Delta(t) &= \Delta(x(t),y(t)) \\ &= -\tfrac{1}{12}x^2(t)y^2(t) - \tfrac{1}{24}(x^4(t)+y^4(t))\end{aligned} \tag{81}$$

The map (80) consists of a standard quadratic dominant part and an additive time-varying disturbance which is state-dependent. Setting $\theta = [x,y]^T = [\hat{x} + a\sin(\omega_1 t), \hat{y} + a\sin(\omega_2 t)]^T$, applying the ES in Fig. 3, we arrive at

$$\begin{aligned}\dot{\tilde{\theta}}(t) &= \tfrac{2k}{a}\begin{bmatrix}\sin(\omega_1 t)\\ \sin(\omega_2 t)\end{bmatrix}\cdot\left[-\tfrac{1}{2}(x^2(t)+y^2(t)) + \Delta(t)\right] \\ &= -k\tilde{\theta}(t) - kL_0(t)\tilde{\theta}(t) - \tfrac{k}{a}S_0(t)\tilde{\theta}^T(t)\tilde{\theta}(t) \\ &\quad -kaS_0(t)S_0^T(t)S_0(t) + \tfrac{2k}{a}S_0(t)\Delta(t)\end{aligned} \tag{82}$$

The application of the averaging or the time-delay approach to the last term $\frac{2k}{a}S_0(t)\Delta(t)$ is not available, as the calculation of the time-derivative of $\Delta(t)$ is complicated.

5. Conclusion

This paper extends the time-delay approach to ES of general nonlinear static maps which are not necessarily requested to be of a quadratic form. Under the premise of a prior knowledge about the upper bounds of the nonlinear map and its gradient and Hessian, the time-delay approach builds a precisely explicit relation between the tuning parameters and the ultimate bound of the estimation error, which suggests a quantitative guideline for the choice of tuning parameters. In the future, expanding the time-delay approach to encompass dynamical maps with low- and high-pass filters in the ES loop is worthy of investigation.

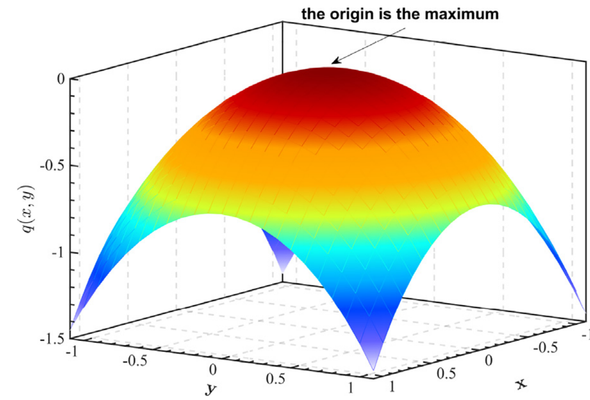


Fig. 5. The shape of the two-variable map.

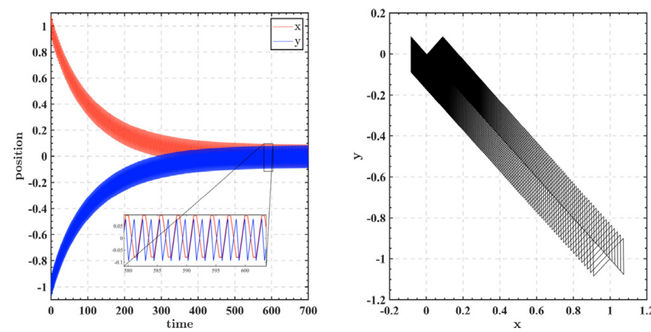


Fig. 6. The trajectory of the ES algorithm.

Acknowledgments

This work was partially supported by National Key R&D Program of China (Grant Nos. 2021YFB1714100, 2018YFA0703803), National Natural Science Foundation of China (Grant No. 62303410, Science Fund Program for Distinguished Young Scholars (Overseas)), Zhejiang Provincial Natural Science Foundation of China (Grant No. LQ23F030014), Israel Science Foundation (Grant No. 673/19) and the ISF-NSFC joint research program (Grant No. 3054/23).

Appendix. Proof of Theorem 2

We first concentrate on the system (71) and define

$$\Phi(t) = \frac{\partial q}{\partial \theta}^T\left(\theta^* + z(t) + kG(t)\right) - \frac{\partial q}{\partial \theta}^T\left(\theta^* + z(t)\right) \tag{A.1}$$

By using differential mean value theorem, and under (45) and (74), we obtain

$$|\Phi(t)| \le \left|\frac{\partial^2 q}{\partial \theta^2}(\theta^* + z(t) + \zeta kG(t))\right| |kG(t)| < q_2(\sigma)\frac{k\Delta_{\mathcal{F}}}{2}\varepsilon \tag{A.2}$$

where $\zeta \in (0,1)$. Refer to Lemma 3.1 on Pages 89–90 of Khalil (2002).

Then the system (71) is expressed as

$$\begin{aligned}\dot{z}(t) &= k\frac{\partial q}{\partial \theta}^T(\theta^* + z(t)) + k\Phi(t)\\ &-\frac{2k}{a}Y_1(t) - kY_2(t) + ka\mathcal{R}(t), \quad t \geq \varepsilon\end{aligned} \tag{A.3}$$

The Lyapunov candidate is selected as

$$V(t) = z^T(t)z(t), \tag{A.4}$$

and its derivative is calculated as

$$\begin{aligned}&\dot{V}(t) + 2\delta V(t) = 2kz^T(t)\frac{\partial q}{\partial \theta}^T(\theta^* + z(t))\\ &+2\delta z^T(t)z(t) + 2kz^T(t)\Phi(t)\\ &-\frac{4k}{a}z^T(t)Y_1(t) - 2kz^T(t)Y_2(t) + 2kaz^T(t)\mathcal{R}(t)\end{aligned} \tag{A.5}$$

We first assume (and then prove) that (72) holds, i.e. $\left|\tilde{\theta}(t)\right| < \sigma$ for all $t \geq 0$.

Under (44) in Assumption 3, we have

$$2kz^T(t)\frac{\partial q}{\partial \theta}^T(\theta^* + z(t)) \leq -2k\mu(\sigma)z^T(t)z(t) \tag{A.6}$$

Then, for any $t \geq \varepsilon$, we employ S-procedure (see Fridman (2014, Chapter 3.2.3)) with the positive parameters λ_1, λ_2, λ_3, λ_4, λ_5 to arrive at

$$\begin{aligned}&\dot{V}(t) + 2\delta V(t) - \frac{\lambda_1}{\varepsilon}G^T(t)G(t)\\ &-\frac{\lambda_2}{\varepsilon}Y_1^T(t)Y_1(t) - \frac{\lambda_3}{\varepsilon}Y_2^T(t)Y_2(t) - \lambda_4 a\mathcal{R}^T(t)\mathcal{R}(t)\\ &+\lambda_5\left(k^2q_2^2(\sigma)G^T(t)G(t) - \Phi^T(t)\Phi(t)\right)\\ &\leq \eta^T(t)\Omega_1\eta(t) < 0\end{aligned} \tag{A.7}$$

where $\eta(t) = \mathrm{col}\{z(t), G(t), Y_1(t), Y_2(t), \Phi(t), \mathcal{R}(t)\}$, and $\Omega_1 < 0$ given in (75) guarantees (A.7) holds.

Note that $\lambda_5\left(k^2q_2^2(\sigma)G^T(t)G(t) - \Phi^T(t)\Phi(t)\right) > 0$ from (A.2). Combining (45) with (74), we get

$$\begin{aligned}&\dot{V}(t) + 2\delta V(t) < \frac{\lambda_1}{\varepsilon}G^T(t)G(t)\\ &+\frac{\lambda_2}{\varepsilon}Y_1^T(t)Y_1(t) + \frac{\lambda_3}{\varepsilon}Y_2^T(t)Y_2(t) + \lambda_4 a\mathcal{R}^T(t)\mathcal{R}(t)\\ &< \frac{\lambda_1\Delta_{\mathcal{F}}^2}{4}\varepsilon + \frac{\lambda_2\bar{S}_0^2q_1^2(\sigma)\Delta_\theta^2}{4}\varepsilon + \frac{\lambda_3\bar{L}_0^2q_2^2(\sigma)\Delta_\theta^2}{4}\varepsilon + \lambda_4 a\Delta_{\mathcal{R}}^2\end{aligned} \tag{A.8}$$

Applying the comparison principle to (A.8), we get

$$\begin{aligned}&V(t) \leq V(\varepsilon)e^{-2\delta(t-\varepsilon)} + \left(1 - e^{-2\delta(t-\varepsilon)}\right)\\ &\times\frac{\left(\lambda_1\Delta_{\mathcal{F}}^2 + \lambda_2\bar{S}_0^2q_1^2(\sigma)\Delta_\theta^2 + \lambda_3\bar{L}_0^2q_2^2(\sigma)\Delta_\theta^2\right)\varepsilon + 4\lambda_4 a\Delta_{\mathcal{R}}^2}{8\delta}\\ &\triangleq U(t), \quad t \geq \varepsilon\end{aligned} \tag{A.9}$$

From (73), we conclude that

$$\begin{aligned}&\left|\tilde{\theta}(t)\right| = \left|\tilde{\theta}(0) + \int_0^t \dot{\tilde{\theta}}(\tau)d\tau\right| < \left|\tilde{\theta}(0)\right| + \Delta_\theta t\\ &\leq \sigma_0 + \Delta_\theta\varepsilon, \quad t \in [0, \varepsilon]\end{aligned} \tag{A.10}$$

The above equation corresponds to the 1st formula in (76) and is ensured by $\Omega_2 < \sigma$ in (75).

Further, from (70), we have

$$\begin{aligned}&\left|\tilde{\theta}(t)\right| \leq |z(t)| + k\,|G(t)| = \sqrt{V(t)} + k\,|G(t)|\\ &< \sqrt{U(t)} + \frac{k\Delta_{\mathcal{F}}}{2}\varepsilon, \quad t \geq \varepsilon\end{aligned} \tag{A.11}$$

which corresponds to the 2nd formula in (76).

Note that $U(t)$ in (A.9) takes the form of $U(t) = \alpha e^{-2\delta(t-\varepsilon)} + \beta$ where $\alpha > 0$ and $\beta > 0$ are constants, hence $U(t)$ is monotonically decreasing. From (A.11), to make sure that $\left|\tilde{\theta}(t)\right| < \sqrt{U(t)} + \frac{k\Delta_{\mathcal{F}}}{2}\varepsilon < \sigma$ for all $t \in [\varepsilon, \infty)$, which is equivalent to the 2nd equation in (76), we need to ensure the condition at the two boundaries such that

$$\begin{cases}\left|\tilde{\theta}(\varepsilon)\right| < \sqrt{U(\varepsilon)} + \frac{k\Delta_{\mathcal{F}}}{2}\varepsilon < \sigma,\\ \lim_{t\to\infty}\left|\tilde{\theta}(t)\right| < \lim_{t\to\infty}\sqrt{U(t)} + \frac{k\Delta_{\mathcal{F}}}{2}\varepsilon < \sigma\end{cases} \tag{A.12}$$

From (A.4), we obtain

$$\begin{aligned}&U(\varepsilon) = V(\varepsilon) = |z(\varepsilon)|^2 \leq \left(\left|\tilde{\theta}(\varepsilon)\right| + k\,|G(\varepsilon)|\right)^2\\ &< \left(\sigma_0 + \varepsilon\Delta_\theta + \frac{k\Delta_{\mathcal{F}}}{2}\varepsilon\right)^2\end{aligned} \tag{A.13}$$

Substituting (A.13) into the 1st row in (A.12), the 1st formula of (A.12) is guaranteed by $\Omega_2 < \sigma$ in (75), whereas the 2nd formula of (A.12) is guaranteed by $\Omega_3 < \left(\sigma - \frac{k\Delta_{\mathcal{F}}}{2}\varepsilon^*\right)^2$ in (75).

We prove next that LMI conditions (75) guarantee that (72) holds via an argument of contradiction (Zhu & Fridman, 2022, Appendix, Page12). Consider first $t \in [0, \varepsilon]$. Since $\left|\tilde{\theta}(0)\right| \leq \sigma_0 < \sigma$ and $\tilde{\theta}(t)$ is continuous in time, (72) holds for small enough $t > 0$. We assume by contradiction that for some $t \in (0, \varepsilon]$ the formula (72) does not hold. Namely, there exists the smallest time instance $t^* \in (0, \varepsilon]$ such that $\left|\tilde{\theta}(t^*)\right| = \sigma$ and $\left|\tilde{\theta}(t)\right| < \sigma$ when $t \in [0, t^*)$. Thus $\left|\tilde{\theta}(t)\right| \leq \sigma$ holds for all $t \in [0, t^*]$ and this leads to the inequality (A.10) in its non-strict version such that $\left|\tilde{\theta}(t)\right| \leq \left|\tilde{\theta}(0)\right| + t\Delta_\theta \leq \sigma_0 + \varepsilon^*\Delta_\theta$ for $0 \leq t \leq t^* \leq \varepsilon \leq \varepsilon^*$. Furthermore, the feasibility of Φ_2 in (75) ensures that $\left|\tilde{\theta}(t^*)\right| \leq \sigma_0 + \varepsilon^*\Delta_\theta < \sigma$. This contradicts to the definition of t^* such that $\left|\tilde{\theta}(t^*)\right| = \sigma$. Hence $\left|\tilde{\theta}(t)\right| < \sigma$ for $t \in [0, \varepsilon]$. Then, we prove (72) for $t \geq \varepsilon$. Since the 1st inequality of (76) derived from (A.10) is strict and holds for $t = \varepsilon$, it infers (72) for sufficiently small $t > \varepsilon$ due to continuity of $\tilde{\theta}(t)$. We assume by contradiction that at some $t > \varepsilon$ the bound (72) does not hold. In other words, there exists the smallest time instance $t^* \in (\varepsilon, \infty)$ such that $\left|\tilde{\theta}(t^*)\right| = \sigma$ and $\left|\tilde{\theta}(t)\right| < \sigma$ when $t \in [\varepsilon, t^*)$. Thus $\left|\tilde{\theta}(t)\right| \leq \sigma$ holds for all $t \in [\varepsilon, t^*]$ and this leads to (A.11) and (A.13) in its non-strict version for $\varepsilon \leq t \leq t^*$. Moreover, the feasibility of Φ_2, Φ_3 in (75) ensures $\left|\tilde{\theta}(t)\right| < \sigma$ in the 2nd equality of (76) for any $t \in [\varepsilon, t^*]$. This contradicts to the definition of t^* such that $\left|\tilde{\theta}(t^*)\right| = \sigma$. Hence $\left|\tilde{\theta}(t)\right| < \sigma$ for $t \geq \varepsilon$.

Finally, we verify the feasibility of (75) for fixed σ and small enough ε^* and a. Note that $\Delta_{\mathcal{F}}$, Δ_θ are of the order of $O\left(\frac{1}{a}\right)$. Choosing $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 1$, clearly $\Omega_2 < 0$ and $\Omega_3 < 0$ in (75) hold for small enough ε^* and $a = (\varepsilon^*)^{\frac{1}{4}}$. Next, we apply Schur complement to Ω_1 of (75), then $\Omega_1 < 0$ is equivalent to

$$\begin{aligned}&\begin{bmatrix}-2(k\mu(\sigma)-\delta)I & 0\\ * & -\left(\frac{\lambda_1}{\varepsilon^*} - \lambda_5 k^2q_2^2(\sigma)\right)I\end{bmatrix} + \frac{\varepsilon^*}{\lambda_2 a^2}\begin{bmatrix}-2kI\\ 0\end{bmatrix}\begin{bmatrix}-2kI & 0\end{bmatrix}\\ &+\frac{\varepsilon^*}{\lambda_3}\begin{bmatrix}-kI\\ 0\end{bmatrix}\begin{bmatrix}-kI & 0\end{bmatrix} + \frac{a}{\lambda_4}\begin{bmatrix}kI\\ 0\end{bmatrix}\begin{bmatrix}kI & 0\end{bmatrix} + \frac{1}{\lambda_5}\begin{bmatrix}kI\\ 0\end{bmatrix}\begin{bmatrix}kI & 0\end{bmatrix} < 0\end{aligned}$$

Choosing $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 1$, $\lambda_5 = \frac{1}{\sqrt{\varepsilon^*}}$, and $a = (\varepsilon^*)^{\frac{1}{4}}$, it is evident that the latter inequality holds for small enough ε^*.

Theorem 2 is proved.

References

Ariyur, Kartik B., & Krstic, Miroslav (2003). *Real-time optimization by extremum-seeking control*. John Wiley & Sons.

Dürr, Hans-Bernd, Stanković, Miloš S., Ebenbauer, Christian, & Johansson, Karl Henrik (2013). Lie bracket approximation of extremum seeking systems. *Automatica*, *49*(6), 1538–1552.

Fridman, Emilia (2014). *Introduction to time-delay systems*. Birkhauser.

Fridman, Emilia, & Shaikhet, Leonid (2016). Delay-induced stability of vector second-order systems via simple Lyapunov functionals. *Automatica*, *74*, 288–296.

Fridman, Emilia, & Zhang, Jin (2020). Averaging of linear systems with almost periodic coefficients: A time-delay approach. *Automatica*, *122*, Article 109287.

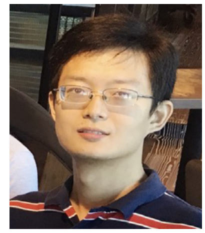
Guay, Martin (2021). Uncertainty estimation in extremum seeking control of unknown static maps. *IEEE Control Systems Letters*, *5*(4), 1115–1120.

Guay, Martin, & Dochain, Denis (2015). A time-varying extremum-seeking control approach. *Automatica*, *51*, 356–363.

Haring, Mark, Fossy, Synne, Silva, Thiago Lima, & Pavlov, Alexey (2022). Nondisturbing extremum seeking control for multi-agent industrial systems. *IEEE Transactions on Automatic Control*, http://dx.doi.org/10.1109/TAC.2022.3153228.

Hazeleger, Leroy, Nesic, Dragan, & van de Wouw, Nathan (2022). Sampled-data extremum-seeking framework for constrained optimization of nonlinear dynamical systems. *Automatica*, *142*, Article 110415.
Khalil, Hassan K. (2002). *Vol. 115*, *Nonlinear systems* (3rd ed.). Patience Hall.
Krilasevic, Suad, & Grammatico, Sergio (2021). Learning generalized Nash equilibria in multi-agent dynamical systems via extremum seeking control. *Automatica*, *133*, Article 109846.
Krstić, Miroslav, & Wang, Hsin-Hsiung (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, *36*(4), 595–601.
Labar, Christophe, Ebenbauer, Christian, & Marconi, Lorenzo (2022). ISS-like properties in Lie-bracket approximations and application to extremum seeking. *Automatica*, *136*, Article 110041.
Liu, Shu-Jun, & Krstic, Miroslav (2012). *Stochastic averaging and stochastic extremum seeking*. Springer Science & Business Media.
Malisoff, Michael, & Krstic, Miroslav (2021). Multivariable extremum seeking with distinct delays using a one-stage sequential predictor. *Automatica*, *129*, Article 109462.
Mele, Adriano, Tommasi, Gianmaria De, & Pironti, Alfredo (2022). Finite-time stabilization of linear systems with unknown control direction via extremum seeking. *IEEE Transactions on Automatic Control*, *67*(10), 5594–5601.
Oliveira, Tiago Roux, Krstic, Miroslav, & Tsubakino, Daisuke (2016). Extremum seeking for static maps with delays. *IEEE Transactions on Automatic Control*, *62*(4), 1911–1926.
Scheinker, Alexander, & Krstić, Miroslav (2017). *Model-free stabilization by extremum seeking*. Springer.
Tan, Ying, Moase, William H., Manzie, Chris, Nešić, Dragan, & Mareels, Iven MY. (2010). Extremum seeking from 1922 to 2010. In *Proceedings of the 29th Chinese control conference* (pp. 14–26). IEEE.
Tan, Ying, Nešić, Dragan, & Mareels, Iven (2006). On non-local stability properties of extremum seeking control. *Automatica*, *42*(6), 889–903.
Yang, Xuefei, & Fridman, Emilia (2023). A robust time-delay approach to continuous-time extremum seeking for multi-variable static map. In *The 62nd IEEE conference on decision and control*. Singapore.
Zhang, Jin, & Fridman, Emilia (2023). Lie-brackets-based averaging of affine systems via a time-delay approach. *Automatica*, *152*, Article 110971.
Zhu, Yang, & Fridman, Emilia (2022). Extremum seeking via a time-delay approach to averaging. *Automatica*, *135*, Article 109965.
Zhu, Yang, & Fridman, Emilia (2023). A time-delay approach for extremum seeking of nonlinear static maps. In *The 12th IFAC symposium on nonlinear control systems*. Canberra, Australia.
Zhu, Yang, Fridman, Emilia, & Oliveira, Tiago Roux (2023). Sampled-data extremum seeking with constant delay: a time-delay approach. *IEEE Transactions on Automatic Control*, *68*, 432–439.
Zhu, Yang, & Krstic, Miroslav (2020). *Delay-adaptive linear control*. Princeton University Press.

Gaofeng Pan was born in 1999. He received his bachelor's degree in Automation from Zhejiang University in July 2021. He is currently pursuing the Ph.D. degree in Control Science and Engineering with Zhejiang University. His research areas include nonlinear systems and extremum seeking.



Yang Zhu received his Bachelor degree from Hangzhou Dianzi University in 2010, and Ph.D. degree from Zhejiang University in 2015. He was a Visiting Graduate Student at University of California, San Diego from 2013 to 2014. During the years 2015–2021, he had served as a Post-doc or Research Fellow at University of Texas at Dallas, Nanyang Technological University, and Tel Aviv University. Since 2021, he has been a Tenure-Track Assistant Professor (in the title of ZJU100 Young Professor) with the College of Control Science and Engineering and the Ningbo Innovation Center of Zhejiang University. He received the Science Fund Program for Distinguished Young Scholars (Overseas) of National Natural Science Foundation of China in 2022. He has served as an Associate Editor in Automatica since 2023. His research focuses on theory of control, optimization, and learning with applications in intelligent transportation systems, and fuel cell systems. Yang is the author of one book with Princeton University Press and a few of research articles with IEEE TAC, Automatica, SICON, SCL, IEEE TITS, IEEE TVT, TR Part C.

Emilia Fridman received the Ph.D. degree from Voronezh State University, USSR, in 1986, in mathematics. Since 1993 she has been at Tel Aviv University, where she is currently Professor of Electrical Engineering Systems. Her research interests include time-delay systems, networked control systems, distributed parameter systems, robust control, singular perturbations and nonlinear control. She has published more than 200 journal articles and two monographs. She serves/served as Associate Editor in Automatica, SIAM Journal on Control and Optimization and IMA Journal of Mathematical Control and Information. She was ranked as a Highly Cited Researcher by Thomson Reuters (Web of Science) in 2014 and the 25th (top 0.03%) in Industrial Engineering & Automation by Stanford University study in 2022. Since 2018, she has been the incumbent for Chana and Heinrich Manderman Chair on System Control at Tel Aviv University. She is IEEE Fellow. In 2021 she was recipient of IFAC (International Federation of Automatic Control) Delay Systems Life Time Achievement Award and of Kadar Award for outstanding research in Tel Aviv University. She is a member of the IFAC Council for 2017–2023. She is IEEE CSS Distinguished Lecturer for 2023–2025. In 2023 her monograph "Introduction to Time-Delay Systems: Analysis and Control" (Birkhauser, 2014) was the winner of IFAC Harold Chestnut Control Engineering Textbook Prize.

Zhengguang Wu was born in 1982. He received the B.S. and M. S. degrees in Mathematics from Zhejiang Normal University, Jinhua, China, in 2004 and 2007, respectively, and the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2011. He now is a professor of Zhejiang university. His research interests include control theory and applications.