2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

A robust time-delay approach to continuous-time extremum seeking for multi-variable static map

Xuefei Yang and Emilia Fridman

***Abstract*— In this article, we introduce a time-delay approach to gradient-based extremum seeking (ES) in the continuous domain for n-dimensional (nD) static quadratic maps. As in the recently introduced (for 2D maps in the continuous domain), we transform the system to the time-delay one (neutral type system). This system is O(ε)-perturbation of the averaged linear ODE system. We further explicitly present the neutral system as the linear ODE, where O(ε)-terms are considered as disturbances with distributed delays of the length of the small parameter ε. Quantitative (for uncertain map) and qualitative (for unknown map) practical stability analyses are provided by employing a variation of constants formula that greatly simplifies the results compared to the previously used Lyapunov-Krasovskii (L-K) method. The new approach also simplifies the conditions and improves the results. Examples from the literature illustrate the efficiency of the new approach, allowing essentially large uncertainty of the Hessian matrix with bounds on ε that are not too small.**

***Index Terms*— Extremum seeking, averaging, time-delay, practical stability.**

I. Introduction

ES a model-free based method for adaptive control which deals with systems where the reference-to-output map is uncertain but is known to have an extremum. Due to its numerous advantages such as its simple principle, low computational complexity and model-free nature, ES control is used in many fields (see [1]-[3]). Following the emergence of the main theoretical breakthrough for an ES system by using averaging and singular perturbations in [4], a great amount of theoretical studies on ES have emerged in the literature, for instance: semi-global and global ES control ([6], [7]), stochastic ES via the stochastic averaging theory ([8], [9]), ES via Lie bracket approximation ([10], [11], [12]) and ES via Newton method ([13], [14]) and ES via hybrid dynamic inclusions ([15]).

The conventional approach to analyze the stability of ES systems is dependent upon the classical averaging theory and singular perturbations (see [16]). However, these methods only provide the qualitative analysis, and cannot suggest quantitative upper bounds on the parameter that preserves the stability. Recently a new constructive time-delay approach to the continuous-time averaging was presented in [17] with efficient and quantitative bounds on the small parameter that ensures the stability. The time-delay approach to averaging was successfully extended to discrete-time systems in [18], and applied for the quantitative stability analysis of continuous-time ES algorithms in [19] and sampled-data ES algorithms in [20] in the case of 2D static maps by constructing appropriate L-K functionals. However, the analysis via L-K method is complicated and the results are conservative, since only small uncertainties in Hessian and initial conditions are available.

In this paper, we develop a robust time-delay approach to ES of nD static maps via practical stability analysis of the averaged system. After transforming the ES dynamics into a time-delay neutral type model as in [19], [20], we further transform it into an averaged ODE perturbed model, and then use the variation of constants formula instead of L-K method to quantitatively analyze the practical stability of the ODE system (and thus of the original ES system). Explicit conditions in terms of simple inequalities are established to guarantee the practical stability of the ES control systems. Through the solution of the constructed inequalities, we find upper bounds on the dither period that ensures the practical stability. Comparatively to the L-K method utilized for neutral type systems in [19], [20], here we adopt the variation of constants formula for the ODE systems. This greatly simplifies the stability analysis process along with the stability conditions, and improve the quantitative bounds as well as the permissible range of the extremum value and the Hessian matrix. Moreover, our approach allows a larger decay rate and a smaller ultimate bound on the estimation error. In addition, for the case that the map is totally unknown, we also provide a simple qualitative analysis by using the variation of constants formula.

Notation: The notation used in this article is fairly standard. The notations $\mathbf{N}_+$, $\mathbf{N}$ and $\mathbf{Z}$ refer to the set of positive integers, nonnegative integers and integers, respectively. The notation $P > 0$ for $P \in \mathbf{R}^{n\times n}$ means that P is symmetric and positive definite. The symmetric elements of the symmetric matrix are denoted by $*$. The notations $|\cdot|$ and $\|\cdot\|$ refer to the usual Euclidean vector norm and the induced matrix 2 norm, respectively.

II. A Time-Delay Approach to ES

A. ES for uncertain map

Consider the multi-variable static map given by

$$y(t) = Q(\theta(t)) = Q^* + \frac{1}{2}[\theta(t) - \theta^*]^{\mathrm{T}} H[\theta(t) - \theta^*], \qquad (1)$$

*This work was supported by in part by the National Natural Science Foundation of China (Grant No 61903102), the Fundamental Research Funds for the Central Universities, the Planning and Budgeting Committee (PBC) Fellowship from the Council for Higher Education in Israel and by Israel Science Foundation (Grant No 673/19).

Xuefei Yang is with Center for Control Theory and Guidance Technology, Harbin Institute of Technology, Harbin, China, and also with the National Key Laboratory of Modeling and Simulation for Complex Systems (e-mail: xfyang1989@163.com, yangxuefei@hit.edu.cn).

Emilia Fridman is with the School of Electrical Engineering, Tel-Aviv University, Israel (e-mail: emilia@tauex.tau.ac.il).

where $y(t) \in \mathbf{R}$ is the measurable output, $\theta(t) \in \mathbf{R}^n$ is the vector input, $Q^* \in \mathbf{R}$ and $\theta^* \in \mathbf{R}^n$ are constants, $H = H^{\mathrm{T}} \in \mathbf{R}^{n\times n}$ is the Hessian matrix which is either positive definite or negative definite. Without loss of generality, we assume that the static map (1) has a minimum value $y(t) = Q^*$ at $\theta(t) = \theta^*$, and then $H > 0$. In the present paper, in order to derive efficient conditions, we assume that:

A1 The extremum point θ^* to be sought is uncertain from a known ball where each of its elements satisfies $\theta_i^*(0) \in [\underline{\theta}_i^*, \bar{\theta}_i^*]$ $(i = 1,\dots,n)$ with $\sum_{i=1}^n (\bar{\theta}_i^* - \underline{\theta}_i^*)^2 = \sigma_0^2$.

A2 The extremum value Q^* is unknown, but it is subject to $|Q^*| \le Q_M^*$ with Q_M^* being known.

A3 The Hessian matrix H is uncertain, $H = \bar{H} + \Delta H$ with $\bar{H} > 0$ being known and $\|\Delta H\| \le \kappa$. Here $\kappa \ge 0$ is a given scalar.

Under **A3**, there exist two positive scalars H_m and H_M such that

$$H_m \le \|H\| \le H_M. \tag{2}$$

Remark 1: In classical ES, the Hessian H, the extremum value Q^* and the extremum point θ^* in (1) are assumed to be unknown, where tuning parameters may be found from simulations only. Here we study a "grey box" with Assumptions **A1-A3** and provide a quantitative analysis. There is a tradeoff between the quantitative analysis with the plant information and the qualitative analysis without the model knowledge.

The gradient-based classical ES algorithm is governed by the following equations:

$$\theta(t) = \hat{\theta}(t) + S(t),\ \dot{\hat{\theta}}(t) = KM(t)y(t), \tag{3}$$

where $\hat{\theta}(t)$ is the real-time estimate of θ^*, $S(t)$ and $M(t)$ are the dither signals satisfying

$$\begin{aligned} S(t) &= [a_1\sin(\omega_1 t),\dots,a_n\sin(\omega_n t)]^{\mathrm{T}},\\ M(t) &= \left[\tfrac{2}{a_1}\sin(\omega_1 t),\dots,\tfrac{2}{a_n}\sin(\omega_n t)\right]^{\mathrm{T}}, \end{aligned} \tag{4}$$

in which $\omega_i \ne \omega_j, i \ne j$ are non-zero, ω_i/ω_j is rational and a_i are real number. The adaptation gain K is chosen as

$$K = \mathrm{diag}\{k_1,k_2,\dots,k_n\},\ k_i < 0,\ i = 1,\dots,n. \tag{5}$$

such that $K\bar{H}$ is Hurwitz (for instance, $K = kI_n$ with a scalar $k < 0$).

Define the estimation error $\tilde{\theta}(t)$ as $\tilde{\theta}(t) = \hat{\theta}(t) - \theta^*$. Then by (3), the estimation error is governed by

$$\begin{aligned} \dot{\tilde{\theta}}(t) = KM(t)\big[&Q^* + \tfrac{1}{2}S^{\mathrm{T}}(t)HS(t) + \tfrac{1}{2}\tilde{\theta}^{\mathrm{T}}(t)H\tilde{\theta}(t)\\ &+ S^{\mathrm{T}}(t)H\tilde{\theta}(t)\big]. \end{aligned} \tag{6}$$

For the stability analysis of the ES control system (6), the classical averaging approach usually resorts to the averaged system via the averaging theorem [16]. To be specific, treating $\tilde{\theta}(t)$ as a “freeze” constant in the averaging analysis and defining $\omega_i = \frac{2\pi l_i}{\varepsilon}, l_i \in \mathbf{N}_+$ $(i = 1,\dots,n)$ satisfying $l_i \ne l_j, i \ne j$, the averaged system of (6) can be derived as [14]

$$\dot{\tilde{\theta}}_{\mathrm{av}}(t) = KH\theta_{\mathrm{av}}(t), \tag{7}$$

which is exponentially stable for small κ since $K\bar{H}$ is Hurwitz. The classical averaging approach leads to a qualitative analysis, and cannot suggest quantitative lower bounds on the dither frequency that guarantee the practical stability as well as the quantitative calculation of the ultimate bound of seeking error. By comparison, we will present an approximation-free method with quantitative bounds on the parameters.

Inspired by [17], [19], we first apply the time-delay approach to averaging of (6). Integrating (6) in $t \ge \varepsilon$ from $t-\varepsilon$ to t, we get

$$\begin{aligned} \tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}\dot{\tilde{\theta}}(\tau)\mathrm{d}\tau = &\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}KM(\tau)Q^*\mathrm{d}\tau\\ &+\tfrac{1}{2\varepsilon}\textstyle\int_{t-\varepsilon}^{t}KM(\tau)S^{\mathrm{T}}(\tau)HS(\tau)\mathrm{d}\tau\\ &+\tfrac{1}{2\varepsilon}\textstyle\int_{t-\varepsilon}^{t}KM(\tau)\tilde{\theta}^{\mathrm{T}}(\tau)H\tilde{\theta}(\tau)\mathrm{d}\tau\\ &+\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}KM(\tau)S^{\mathrm{T}}(\tau)H\tilde{\theta}(\tau)\mathrm{d}\tau,\ t \ge \varepsilon. \end{aligned} \tag{8}$$

In the remainder of this paper, we define $x \pm y \triangleq x + y - y$. For the first term on the right-hand side of (8), we have

$$\begin{aligned} &\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}KM(\tau)Q^*\mathrm{d}\tau\\ &= \tfrac{1}{\varepsilon}Q^*K\,\mathrm{col}\left\{\tfrac{2}{a_i}\textstyle\int_{t-\varepsilon}^{t}\sin\left(\tfrac{2\pi l_i}{\varepsilon}\tau\right)\mathrm{d}\tau\right\}_{i=1}^{n} = 0, \end{aligned} \tag{9}$$

where we have used

$$\textstyle\int_{t-\varepsilon}^{t}\sin\left(\tfrac{2\pi l_i}{\varepsilon}\tau\right)\mathrm{d}\tau = 0,\ i = 1,\dots,n. \tag{10}$$

For the second term on the right-hand side of (8), we have

$$\begin{aligned} &\tfrac{1}{2\varepsilon}\textstyle\int_{t-\varepsilon}^{t}KM(\tau)S^{\mathrm{T}}(\tau)HS(\tau)\mathrm{d}\tau\\ &= \tfrac{1}{\varepsilon}K\,\mathrm{col}\Bigg\{\sum_{i=1}^{n}\sum_{j=1}^{n}\tfrac{a_ia_jh_{ij}}{a_k}\textstyle\int_{t-\varepsilon}^{t}\sin\left(\tfrac{2\pi l_i}{\varepsilon}\tau\right)\\ &\quad\times\sin\left(\tfrac{2\pi l_j}{\varepsilon}\tau\right)\sin\left(\tfrac{2\pi l_k}{\varepsilon}\tau\right)\mathrm{d}\tau\Bigg\}_{k=1}^{n} = 0, \end{aligned} \tag{11}$$

where we have utilized

$$\textstyle\int_{t-\varepsilon}^{t}\sin\left(\tfrac{2\pi l_i}{\varepsilon}\tau\right)\sin\left(\tfrac{2\pi l_j}{\varepsilon}\tau\right)\sin\left(\tfrac{2\pi l_k}{\varepsilon}\tau\right)\mathrm{d}\tau = 0.$$

For the third term on the right-hand side of (8), we have

$$\begin{aligned} &\tfrac{1}{2\varepsilon}\textstyle\int_{t-\varepsilon}^{t}KM(\tau)\tilde{\theta}^{\mathrm{T}}(\tau)H\tilde{\theta}(\tau)\mathrm{d}\tau\\ &= \tfrac{1}{2\varepsilon}\textstyle\int_{t-\varepsilon}^{t}KM(\tau)[\tilde{\theta}^{\mathrm{T}}(\tau)H\tilde{\theta}(\tau) \pm \tilde{\theta}^{\mathrm{T}}(t)H\tilde{\theta}(t)]\mathrm{d}\tau\\ &= \tfrac{1}{2\varepsilon}\tilde{\theta}^{\mathrm{T}}(t)H\tilde{\theta}(t)K\textstyle\int_{t-\varepsilon}^{t}M(\tau)\mathrm{d}\tau\\ &\quad-\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}\int_{\tau}^{t}KM(\tau)\tilde{\theta}^{\mathrm{T}}(s)H\dot{\tilde{\theta}}(s)\mathrm{d}s\mathrm{d}\tau\\ &= -\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}\int_{\tau}^{t}KM(\tau)\tilde{\theta}^{\mathrm{T}}(s)H\dot{\tilde{\theta}}(s)\mathrm{d}s\mathrm{d}\tau, \end{aligned} \tag{12}$$

where we have employed $\int_{t-\varepsilon}^{t}M(\tau)\mathrm{d}\tau = 0$ via (10). For the fourth term on the right-hand side of (8), we have

$$\begin{aligned} &\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}KM(\tau)S^{\mathrm{T}}(\tau)H\tilde{\theta}(\tau)\mathrm{d}\tau\\ &= \tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}KM(\tau)S^{\mathrm{T}}(\tau)H[\tilde{\theta}(\tau) \pm \tilde{\theta}(t)]\mathrm{d}\tau\\ &= \tfrac{1}{\varepsilon}K\textstyle\int_{t-\varepsilon}^{t}M(\tau)S^{\mathrm{T}}(\tau)\mathrm{d}\tau H\tilde{\theta}(t)\\ &\quad-\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}\int_{\tau}^{t}KM(\tau)S^{\mathrm{T}}(\tau)H\dot{\tilde{\theta}}(s)\mathrm{d}s\mathrm{d}\tau\\ &= KH\tilde{\theta}(t) - \tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}\int_{\tau}^{t}KM(\tau)S^{\mathrm{T}}(\tau)H\dot{\tilde{\theta}}(s)\mathrm{d}s\mathrm{d}\tau, \end{aligned} \tag{13}$$

where we have utilized $\int_{t-\varepsilon}^{t}M(\tau)S^{\mathrm{T}}(\tau)\mathrm{d}\tau = \varepsilon I_n$, since

$$\textstyle\int_{t-\varepsilon}^{t}\tfrac{2a_i}{a_j}\sin\left(\tfrac{2\pi l_i}{\varepsilon}\tau\right)\sin\left(\tfrac{2\pi l_j}{\varepsilon}\tau\right)\mathrm{d}\tau = \begin{cases}\varepsilon, & i = j,\\ 0, & i \ne j.\end{cases}$$

For the left-hand side of (8), we have

$$\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}\dot{\tilde{\theta}}(\tau)\mathrm{d}\tau = \tfrac{\mathrm{d}}{\mathrm{d}t}[\tilde{\theta}(t) - G(t)], \tag{14}$$

where

$$G(t) = \tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^{t}(\tau - t + \varepsilon)\dot{\tilde{\theta}}(\tau)\mathrm{d}\tau. \tag{15}$$

Finally, employing (9), (11)-(14), system (8) can be transformed to

$$\frac{\mathrm{d}}{\mathrm{d}t}[\tilde{\theta}(t)-G(t)] = KH\tilde{\theta}(t)-Y_1(t)-Y_2(t),\ t\geq\varepsilon, \quad (16)$$

where

$$\begin{aligned} Y_1(t) &= \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau}^{t}KM(\tau)\tilde{\theta}^{\mathrm{T}}(s)H\dot{\tilde{\theta}}(s)\mathrm{d}s\mathrm{d}\tau, \\ Y_2(t) &= \tfrac{1}{\varepsilon}\int_{t-\varepsilon}^{t}\int_{\tau}^{t}KM(\tau)S^{\mathrm{T}}(\tau)H\dot{\tilde{\theta}}(s)\mathrm{d}s\mathrm{d}\tau, \end{aligned} \quad (17)$$

whereas $\dot{\tilde{\theta}}(s)$ is defined by the right-hand side of (6). Clearly, the solution $\tilde{\theta}(t)$ of system (6) is also a solution of system (16). Thus, the practical stability of the original non-delayed system (6) can be guaranteed by the practical stability of the time-delay system (16), which is a neutral type system with the state $\tilde{\theta}$, as derived in [19] for 2D maps.

In this paper, for simplifying the stability analysis, we further set

$$z(t) = \tilde{\theta}(t)-G(t). \quad (18)$$

Then system (16) can be rewritten as

$$\dot{z}(t) = KHz(t)+KHG(t)-Y_1(t)-Y_2(t),\ t\geq\varepsilon. \quad (19)$$

Comparatively to the averaged system (7), system (19) has the additional terms $G(t)$, $Y_1(t)$ and $Y_2(t)$ that are of the order of $\mathrm{O}(\varepsilon)$ provided $\tilde{\theta}(s)$ and $\dot{\tilde{\theta}}(s)$ (and thus $z(t)$) are of the order of $\mathrm{O}(1)$. Hence, for small $\varepsilon>0$ system (19) can be regarded as a perturbation of system (7).

Differently from [19], we will analyze (19) as ODE w.r.t. z (and not as neutral type w.r.t. $\tilde{\theta}$) with delayed disturbance-like $\mathrm{O}(\varepsilon)$-terms G,Y_1,Y_2 that depend on the solutions of (6). The resulting bound on $|z|$ will lead to the bound on $\tilde{\theta}$: $|\tilde{\theta}|\leq|z|+|G|$. The bound on z will be found by utilizing the variation of constants formula compared to L-K method employed in [19].

Theorem 1: Let **A1-A3** be satisfied. Consider the closed-loop system (6) with the initial condition $|\tilde{\theta}(0)|\leq\sigma_0$. Given tuning parameters k_i, a_i $(i=1,\ldots,n)$ and δ, let matrix P $(I_n\leq P\leq pI_n)$ with a scalar $p\geq 1$ and scalar $\zeta>0$ satisfy the following LMI:

$$\begin{aligned} \Phi_1 &= \begin{bmatrix} \Phi_{11} & PK \\ * & -\zeta I_n \end{bmatrix} < 0, \\ \Phi_{11} &= \bar{H}^{\mathrm{T}}K^{\mathrm{T}}P+PK\bar{H}+2\delta P+\zeta\kappa^2 I_n. \end{aligned} \quad (20)$$

Given $\sigma>\sigma_0>0$, let there exits $\varepsilon^*>0$ that satisfy

$$\Phi_2 = p\left(\sigma_0+\tfrac{\varepsilon^*\Delta[2(\Delta_1+\Delta_2+\Delta_3)+3\delta]}{2\delta}\right)^2-\left(\sigma-\tfrac{\varepsilon^*\Delta}{2}\right)^2<0, \quad (21)$$

where

$$\begin{aligned} \Delta &= \left[Q_M^*+\tfrac{H_M}{2}\left(\sigma+\sqrt{\textstyle\sum_{i=1}^{n}a_i^2}\right)^2\right]\sqrt{\textstyle\sum_{i=1}^{n}\tfrac{4k_i^2}{a_i^2}}, \\ \Delta_1 &= \tfrac{H_M\max_{i\in\mathbf{I}[1,n]}|k_i|}{2},\ \Delta_2=\tfrac{\sigma H_M}{2}\sqrt{\textstyle\sum_{i=1}^{n}\tfrac{4k_i^2}{a_i^2}}, \\ \Delta_3 &= \tfrac{H_M}{2}\sqrt{\textstyle\sum_{i=1}^{n}\tfrac{4k_i^2}{a_i^2}}\sqrt{\textstyle\sum_{i=1}^{n}a_i^2}. \end{aligned} \quad (22)$$

Then for all $\varepsilon\in(0,\varepsilon^*]$, the solution of the estimation error system (6) satisfies

$$\begin{aligned} &|\tilde{\theta}(t)|<|\tilde{\theta}(0)|+\varepsilon\Delta<\sigma,\ t\in[0,\varepsilon], \\ &|\tilde{\theta}(t)|<\sqrt{p}\mathrm{e}^{-\delta(t-\varepsilon)}\left(|\tilde{\theta}(0)|+\tfrac{3\varepsilon\Delta}{2}\right) \\ &\qquad+\tfrac{\varepsilon\Delta[2(\Delta_1+\Delta_2+\Delta_3)\sqrt{p}+\delta]}{2\delta}<\sigma,\ t\geq\varepsilon. \end{aligned} \quad (23)$$

Moreover, for all $\varepsilon\in(0,\varepsilon^*]$ and all initial conditions $|\tilde{\theta}(0)|\leq\sigma_0$, the ball

$$\Theta = \left\{\tilde{\theta}\in\mathbf{R}:|\tilde{\theta}(t)|<\tfrac{\varepsilon\Delta[2(\Delta_1+\Delta_2+\Delta_3)\sqrt{p}+\delta]}{2\delta}\right\} \quad (24)$$

is exponential attractive with a decay rate δ.

Proof: See Appendix A1. ■

Remark 2: Given any σ_0 and $\sigma^2>p\sigma_0^2$, inequality ($\Phi_2<0$ in (21)) is always feasible for small enough ε^*. Therefore, the result is semi-global. For $\Phi_1<0$ in (20), since $K\bar{H}$ is Hurwitz, there exists a $n\times n$ matrix $P>0$ such that for small enough $\delta>0$, the following inequality holds: $\Psi=\bar{H}^{\mathrm{T}}K^{\mathrm{T}}P+PK\bar{H}+2\delta P<0$. We choose $\zeta=1/\kappa$. Applying the Schur complement to $\Phi_1<0$, we have $\Psi+\kappa(I_n+PKKP)<0$, which always holds for small enough $\kappa>0$ since $\Psi<0$. For $P>0$, there exist positive scalars p_1 and p_2 such that

$$p_1 I_n\leq P\leq p_2 I_n. \quad (25)$$

If $p_1\neq 1$, we can rewrite (25) as $I_n\leq\frac{1}{p_1}P\leq\frac{p_2}{p_1}I_n$, which is in the form of $I_n\leq P\leq pI_n$ by setting $P=P/p_1$ and $p=p_2/p_1$. Furthermore, $\Phi_i<0$ $(i=1,2)$ hold with the modified $\{P,p\}$ as well as the bound in (24).

Remark 3: We give a brief discussion about the effect of free parameters on the performance of ES system. For simplicity, let $K=kI_n$ with $k<0$ being a given scalar. Then from (22) we know that Δ and Δ_i $(i=1,2,3)$ are of the order of $\mathrm{O}(|k|)$ as well as the decay rate δ since $\delta=|k|\lambda_{\min}(H)$. Thus

$$\vartheta_1\triangleq\frac{2\sqrt{p}\Delta(\Delta_1+\Delta_2+\Delta_3)+\left(3\sqrt{p}+1\right)\Delta\delta}{2\delta}$$

is of the order of $\mathrm{O}(|k|)$. Note from (21) that $\varepsilon^*<\frac{1}{\vartheta_1}\left(\sigma-\sqrt{p}\sigma_0\right)$, which implies that for given $\sigma>\sigma_0>0$, ε^* is of the order of $\mathrm{O}(1/|k|)$. Therefore, the decay rate δ increases as $|k|$ increases, while ε^* decreases as $|k|$ increases. So we can adjust the gain $K=kI_n$ to balance the decay rate δ and ε^*. In addition, we let $\vartheta_2\triangleq\frac{\Delta[2(\Delta_1+\Delta_2+\Delta_3)\sqrt{p}+\delta]}{2\delta}$. Then the ball in (24) can be rewritten as

$$\Theta = \left\{\tilde{\theta}\in\mathbf{R}:|\tilde{\theta}(t)|<\varepsilon\vartheta_2\right\}. \quad (26)$$

Note from (22) that ϑ_2 is an increasing function of σ, thus, for given σ_0, δ, ε, a_i and k_i $(i=1,\ldots,n)$, we can solve the inequality (21) to find the smallest σ, and then substitute it into (26) to get the bound. Moreover, if $\varepsilon\vartheta_2<\sigma_0-\beta$ with some $\beta\in(0,\sigma_0)$, we can reset $\sigma_0=\varepsilon\vartheta_2+\beta$ and repeat the above process to obtain a smaller ultimate bound (UB). Obviously, the lower bound of UB in theory is $\varepsilon\vartheta_2$ with $\sigma=0$.

Remark 4: Compared with the results in [19], Theorem 1 presents much simpler proof and LMI-based conditions, which allow us to get larger decay rate and period of the dither signal. Moreover, it is observed from (24) that the ultimate bound on the estimation error is of the order of $\mathrm{O}(\varepsilon)$ provided that a_i, k_i $(i=1,\ldots,n)$ are of the order of $\mathrm{O}(1)$ leading to δ of the order of $\mathrm{O}(1)$. This is smaller than $\mathrm{O}(\sqrt{\varepsilon})$ achieved in [19]. In addition, due to the complexity of the LMIs in the vertices when the Hessian H is not known, the work [19] did not go into details to discuss the uncertainty case. As a comparison, by using the established time-delay approach, we can easily solve the uncertainty case.

Next we consider a special case with the Hessian H being diagonal, namely, $H=\operatorname{diag}\{h_1,h_2,\ldots,h_n\}$ with $h_i>0$ $(i=1,\ldots,n)$. We also assume that H is unknown, but satisfies (2). In this case, instead of utilizing the Lyapunov method to find the upper bound of the fundamental matrix e^{KHt}, we can directly compute that

$$\left\|\mathrm{e}^{KHt}\right\| \le \mathrm{e}^{-H_m \min_{i\in\mathbf{I}[1,n]}|k_i| t} \triangleq \mathrm{e}^{-\delta t},\ \forall t\ge 0.$$

This can lead to a simpler analysis and more concise result as shown in the following corollary.

Corollary 1: Let **A1-A2** be satisfied and the diagonal Hessian H be unknown but satisfy (2). Consider the closed-loop system (6) with the initial condition $|\tilde{\theta}(0)|\le\sigma_0$. Given tuning parameters k_i, a_i $(i=1,\ldots,n)$ and $\sigma>\sigma_0>0$, let there exits $\varepsilon^*>0$ that satisfy

$$\Phi=\sigma_0+\tfrac{\varepsilon^*\Delta[(\Delta_1+\Delta_2+\Delta_3)+2\delta]}{\delta}<\sigma,$$

where Δ and Δ_i $(i=1,2,3)$ are given by (22). Then for all $\varepsilon\in(0,\varepsilon^*]$, the solution of the estimation error system (6) satisfies

$$\begin{aligned}&|\tilde{\theta}(t)|<|\tilde{\theta}(0)|+\varepsilon\Delta<\sigma,\ t\in[0,\varepsilon],\\ &|\tilde{\theta}(t)|<\mathrm{e}^{-\delta(t-\varepsilon)}\left(|\tilde{\theta}(0)|+\tfrac{3\varepsilon\Delta}{2}\right)\\ &\qquad+\tfrac{\varepsilon\Delta[2(\Delta_1+\Delta_2+\Delta_3)+\delta]}{2\delta}<\sigma,\ t\ge\varepsilon.\end{aligned}$$

Moreover, for all $\varepsilon\in(0,\varepsilon^*]$ and all initial conditions $|\tilde{\theta}(0)|\le\sigma_0$, the ball

$$\Theta=\left\{\tilde{\theta}\in\mathbf{R}:|\tilde{\theta}(t)|<\tfrac{\varepsilon\Delta[2(\Delta_1+\Delta_2+\Delta_3)+\delta]}{2\delta}\right\}$$

is exponential attractive with a decay rate $\delta=H_m\min_{i\in\mathbf{I}[1,n]}|k_i|$.

B. ES for unknown map

In this section, we consider a more general case that the map $Q(t)$ is totally unknown without the assumptions of **A1-A3** as needed in the previous section. Without loss of generality, we still assume $H>0$ and consider the gradient-based ES algorithm (3) with $\hat{\theta}(0)\in\mathbf{R}^n$, the adaptation gain K is chosen as (5) such that KH is Hurwitz. The estimation error of $\hat{\theta}(t)$ on θ^* satisfies (6). By the arguments similar to (8)-(18), we can obtain (19), which follows from (6). The stability analysis procedure for this case follows the same steps as Theorem 1.

Consider the corresponding to (19) homogeneous system:

$$\dot{z}(t)=KHz(t). \tag{27}$$

Since KH is Hurwitz, then by Theorem 4.11 in [16], there exist two scalars $\alpha,\beta>0$ such that the state transition matrix of (27) satisfies

$$\left\|\mathrm{e}^{KHt}\right\|\leqslant\alpha\mathrm{e}^{-\beta t},\ \forall t\geqslant 0. \tag{28}$$

By the same stability analysis procedure as in Theorem 1 and (28), we obtain

$$\begin{aligned}|z(t)|&<\left\|\mathrm{e}^{KH(t-\varepsilon)}\right\||z(\varepsilon)|\\&\quad+\varepsilon\Delta(\Delta_1+\Delta_2+\Delta_3)\int_\varepsilon^t\left\|\mathrm{e}^{KH(t-s)}\right\|\mathrm{d}s\\&<\alpha\mathrm{e}^{-\beta(t-\varepsilon)}\left(|\tilde{\theta}(0)|+\tfrac{3\varepsilon\Delta}{2}\right)+\tfrac{\alpha\varepsilon\Delta(\Delta_1+\Delta_2+\Delta_3)}{\beta},\ t\ge\varepsilon,\end{aligned}$$

where

$$\begin{aligned}&\Delta=\left[|Q^*|+\tfrac{\|H\|}{2}\left(\sigma+\sqrt{\textstyle\sum_{i=1}^n a_i^2}\right)^2\right]\sqrt{\textstyle\sum_{i=1}^n\tfrac{4k_i^2}{a_i^2}},\\&\Delta_1=\tfrac{\|H\|\max_{i=1,\ldots,n}|k_i|}{2},\ \Delta_2=\tfrac{\sigma\|H\|}{2}\sqrt{\textstyle\sum_{i=1}^n\tfrac{4k_i^2}{a_i^2}},\\&\Delta_3=\tfrac{\|H\|}{2}\sqrt{\textstyle\sum_{i=1}^n\tfrac{4k_i^2}{a_i^2}}\sqrt{\textstyle\sum_{i=1}^n a_i^2},\end{aligned}\tag{29}$$

by which and (18), we finally get

$$\begin{aligned}\left|\tilde{\theta}(t)\right|&\leqslant|z(t)|+|G(t)|\\&<\alpha\mathrm{e}^{-\beta(t-\varepsilon)}\left(|\tilde{\theta}(0)|+\tfrac{3\varepsilon\Delta}{2}\right)\\&\quad+\tfrac{\varepsilon\alpha\Delta(\Delta_1+\Delta_2+\Delta_3)}{\beta}+\tfrac{\varepsilon\Delta}{2},\ t\ge\varepsilon.\end{aligned}$$

Based on the above analysis, the semi-global practical stability of the closed-loop system (6) can be stated in the following proposition:

Proposition 1: For any σ and σ_0 satisfying $\sigma>\alpha\sigma_0>0$, consider the closed-loop system (6) with the initial condition $|\tilde{\theta}(0)|\le\sigma_0$. Given tuning parameters k_i and a_i $(i=1,\ldots,n)$, there exists $\varepsilon^*>0$ such that for all $\varepsilon\in(0,\varepsilon^*]$, the solution of the estimation error system (6) satisfies

$$\begin{aligned}&|\tilde{\theta}(t)|<|\tilde{\theta}(0)|+\varepsilon\Delta<\sigma,\ t\in[0,\varepsilon],\\&|\tilde{\theta}(t)|<\alpha\mathrm{e}^{-\beta(t-\varepsilon)}\left(|\tilde{\theta}(0)|+\tfrac{3\varepsilon\Delta}{2}\right)\\&\qquad+\tfrac{\varepsilon\alpha\Delta(\Delta_1+\Delta_2+\Delta_3)}{\beta}+\tfrac{\varepsilon\Delta}{2}<\sigma,\ t\ge\varepsilon\end{aligned}$$

with some positive α and β, where Δ, Δ_i, $i=1,2,3$ are given by (29). Furthermore,

$$\limsup_{t\to\infty}\left|\tilde{\theta}(t)\right|=\mathscr{O}(\varepsilon).$$

C. Examples

1) Vector systems: $n=2$: Consider an autonomous vehicle in an environment without GPS orientation [5]. The goal is to reach the location of the stationary minimum of a measurable function

$$\begin{aligned}&J(x(t),y(t))=Q^*+\tfrac{1}{2}\begin{bmatrix}x(t) & y(t)\end{bmatrix}H\begin{bmatrix}x(t)\\ y(t)\end{bmatrix}\\&=x^2(t)+y^2(t),\end{aligned}$$

where

$$Q^*=0,\ H=\begin{bmatrix}2&0\\0&2\end{bmatrix}.$$

We employ the classical ES

$$\begin{aligned}&x(t)=\hat{x}(t)+a_1\sin(\omega_1 t),\ y(t)=\hat{y}(t)+a_2\sin(\omega_2 t),\\&\dot{\hat{x}}(t)=\tfrac{2k_1}{a_1}\sin(\omega_1 t)J(t),\ \dot{\hat{y}}(t)=\tfrac{2k_2}{a_2}\sin(\omega_2 t)J(t)\end{aligned}$$

with $k_1=k_2=-0.01$, $a_1=a_2=0.2$. The solutions are shown in Table I. It follows that Corollary 1 allows larger decay rate δ and much larger upper bound ε^* than those in [19].

Moreover, when the upper bound ε^* shares the same value, our results allow much larger uncertainties in initial condition σ_0 than those in [19]. Finally, we make a comparison for the ultimate bound under the same value of ε. The solutions are shown in Table II. It follows that the values of UB obtained by Corollary 1 are much smaller that those in [19].

TABLE I
COMPARISON OF ε^* IN VECTOR SYSTEMS: $n=2$

ES: sine wave	σ_0	σ	δ	ε^*
[19]	$\sqrt{2}$	$2\sqrt{2}$	0.01	0.017
Corollary 1	$\sqrt{2}$	$2\sqrt{2}$	0.02	0.042
Corollary 1	2.55	4	0.02	0.017

TABLE II
COMPARISON OF UB IN VECTOR SYSTEMS: $n=2$

ES: sine wave	σ_0	σ	δ	ε	UB
[19]	$\sqrt{2}$	$2\sqrt{2}$	0.01	0.017	1.9
Corollary 1	2.55	4	0.02	0.017	1.4e^{-3}

2) Vector systems: $n=6$: Consider the quadratic function (1) with [10]

$$Q^*=0,\ \theta^*=[1,1,-1,-1,-1,1]^{\mathrm{T}},\ H=\mathrm{diag}\{1,1,1,1,1,3\}.$$

If Q^* and H are unknown, but satisfy **A2** and (2) we consider

$$Q^*_M=0.5,\ H_m=0.8,\ H_M=3.2.$$

We select the tuning parameters of the gradient-based ES as $k_i=-0.05$, $a_i=1$ $(i=1,\dots,6)$. Both the results for known and uncertain Q are shown in Table III, illustrating the efficiency of our method.

TABLE III
VECTOR SYSTEMS: $n=6$

ES: sine wave	σ_0	σ	δ	ε^*	UB
Uncertainty-free case	1	2	0.150	1.0e^{-2}	0.315
Uncertainty case	1	2	0.025	1.4e^{-3}	0.382

III. CONCLUSION

This paper developed a time-delay approach to ES in the continuous domain which offers a simpler and more efficient stability analysis method based on simple inequalities. Explicit conditions in terms of inequalities were established to guarantee the practical stability of the ES control systems by employing the variation of constants formula to the perturbed averaged system. In comparison to the L-K method, the newly established method not only simplifies the stability analysis of ES but also improves its results. For example, it enables us to obtain larger decay rates, periods of dither signals, and uncertainties of the map. Furthermore, for the case that the map is totally unknown, our approach also provides a simplified qualitative practical stability analysis. Further research topics could include ES for dynamic maps and non-quadratic maps.

APPENDIX: PROOF OF THEOREM 1

Assume that

$$|\tilde{\theta}(t)|<\sigma,\ \forall t\geq 0. \tag{30}$$

Note from (1)-(6) and (30) that

$$\begin{aligned}|y(t)|&=\left|Q^*+\tfrac{1}{2}(\tilde{\theta}(t)+S(t))^{\mathrm{T}}H(\tilde{\theta}(t)+S(t))\right|\\&<Q^*_M+\tfrac{H_M}{2}\left(\sigma+\sqrt{\textstyle\sum_{i=1}^n a_i^2}\right)^2,\ t\geq 0,\\ \left|\dot{\tilde{\theta}}(t)\right|&=|KM(t)y(t)|<\Delta,\ t\geq 0,\\ |\tilde{\theta}(t)|&=\left|\tilde{\theta}(0)+\textstyle\int_0^t\dot{\tilde{\theta}}(s)\mathrm{d}s\right|<|\tilde{\theta}(0)|+\varepsilon\Delta,\ t\in[0,\varepsilon]\end{aligned} \tag{31}$$

with Δ given by (22). The first inequality in (23) follows from the third inequality in (31) since $\Phi_2<0$ in (21) implies that $\sigma_0+\varepsilon^*\Delta<\sigma$, $\forall\varepsilon\in(0,\varepsilon^*]$. Next we consider the case with $t\geq\varepsilon$.

To make the second inequality in (23) hold, we use the variation of constants formula for (19) to obtain

$$\begin{aligned}z(t)&=\mathrm{e}^{KH(t-\varepsilon)}z(\varepsilon)\\&+\textstyle\int_\varepsilon^t\mathrm{e}^{KH(t-s)}[KHG(s)-Y_1(s)-Y_2(s)]\mathrm{d}s,\ t\geq\varepsilon.\end{aligned} \tag{32}$$

From (15) and (31) we have

$$\begin{aligned}|G(t)|&=\left|\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^t(\tau-t+\varepsilon)\dot{\tilde{\theta}}(\tau)\mathrm{d}\tau\right|\\&\leq\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^t\left|(\tau-t+\varepsilon)\dot{\tilde{\theta}}(\tau)\right|\mathrm{d}\tau\\&<\tfrac{1}{\varepsilon}\Delta\textstyle\int_{t-\varepsilon}^t(\tau-t+\varepsilon)\mathrm{d}\tau=\tfrac{\varepsilon\Delta}{2},\end{aligned} \tag{33}$$

and

$$|KHG(t)|\leq\|K\|\,\|H\|\,|G(t)|<\frac{\varepsilon\Delta\cdot H_M\max_{i\in\mathbf{I}[1,n]}|k_i|}{2}=\varepsilon\Delta\cdot\Delta_1 \tag{34}$$

with Δ_1 given by (22). From (17) and (31) we have

$$\begin{aligned}|Y_1(t)|&=\left|\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^t\int_\tau^t KM(\tau)\tilde{\theta}^{\mathrm{T}}(s)H\dot{\tilde{\theta}}(s)\mathrm{d}s\mathrm{d}\tau\right|\\&\leq\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^t\int_\tau^t|KM(\tau)|\left|\tilde{\theta}^{\mathrm{T}}(s)\right|\|H\|\left|\dot{\tilde{\theta}}(s)\right|\mathrm{d}s\mathrm{d}\tau\\&<\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^t\int_\tau^t\sqrt{\sum_{i=1}^n\frac{4k_i^2}{a_i^2}}\sigma H_M\Delta\mathrm{d}s\mathrm{d}\tau\\&=\tfrac{\varepsilon}{2}\sqrt{\textstyle\sum_{i=1}^n\frac{4k_i^2}{a_i^2}}\sigma H_M\Delta=\varepsilon\Delta\cdot\Delta_2,\end{aligned} \tag{35}$$

and

$$\begin{aligned}|Y_2(t)|&=\left|\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^t\int_\tau^t KM(\tau)S^{\mathrm{T}}(\tau)H\dot{\tilde{\theta}}(s)\mathrm{d}s\mathrm{d}\tau\right|\\&\leq\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^t\int_\tau^t|KM(\tau)|\left|S^{\mathrm{T}}(\tau)\right|\|H\|\left|\dot{\tilde{\theta}}(s)\right|\mathrm{d}s\mathrm{d}\tau\\&<\tfrac{1}{\varepsilon}\textstyle\int_{t-\varepsilon}^t\int_\tau^t\sqrt{\sum_{i=1}^n\frac{4k_i^2}{a_i^2}}\sqrt{\sum_{i=1}^n a_i^2}H_M\Delta\mathrm{d}s\mathrm{d}\tau\\&=\tfrac{\varepsilon}{2}\sqrt{\textstyle\sum_{i=1}^n\frac{4k_i^2}{a_i^2}}\sqrt{\textstyle\sum_{i=1}^n a_i^2}H_M\Delta=\varepsilon\Delta\cdot\Delta_3,\end{aligned} \tag{36}$$

where Δ_2 and Δ_3 are given by (22). Via (32) and (34)-(36), we obtain

$$\begin{aligned}|z(t)|&\leq\left\|\mathrm{e}^{KH(t-\varepsilon)}\right\||z(\varepsilon)|\\&+\textstyle\int_\varepsilon^t\left\|\mathrm{e}^{KH(t-s)}\right\|[|KHG(s)|+|Y_1(s)|+|Y_2(s)|]\mathrm{d}s\\&<\left\|\mathrm{e}^{KH(t-\varepsilon)}\right\||z(\varepsilon)|\\&+\varepsilon\Delta(\Delta_1+\Delta_2+\Delta_3)\textstyle\int_\varepsilon^t\left\|\mathrm{e}^{KH(t-s)}\right\|\mathrm{d}s,\ t\geq\varepsilon.\end{aligned} \tag{37}$$

In order to derive a bound on e^{KHt}, consider the nominal system

$$\dot{z}(t) = KHz(t) = K(\bar{H} + \Delta H)z(t),\ t \geq 0, \tag{38}$$

where we noted **A3**. Choose the Lyapunov function $V(t) = z^{\mathrm{T}}(t)Pz(t)$ with P satisfying $I_n \leq P \leq pI_n$. Then

$$\begin{aligned}\dot{V}(t) + 2\delta V(t) = 2z^{\mathrm{T}}(t)P[K(\bar{H} + \Delta H)]z(t) \\ + 2\delta z^{\mathrm{T}}(t)Pz(t).\end{aligned} \tag{39}$$

To compensate $\Delta Hz(t)$ in (39) we apply S-procedure, we add to $\dot{V}(t) + 2\delta V(t)$ the left hand part of $\zeta(\kappa^2|z(t)|^2 - |\Delta Hz(t)|^2) \geq 0$ with some $\zeta > 0$. Then, we have

$$\begin{aligned}&\dot{V}(t) + 2\delta V(t) \leq 2z^{\mathrm{T}}(t)P[K(\bar{H} + \Delta H)]z(t) \\ &\quad + 2\delta z^{\mathrm{T}}(t)Pz(t) + \zeta\left(\kappa^2|z(t)|^2 - |\Delta Hz(t)|^2\right) \\ &= \xi^{\mathrm{T}}(t)\Phi_1\xi(t),\end{aligned}$$

where $\xi^{\mathrm{T}}(t) = [z^{\mathrm{T}}(t), z^{\mathrm{T}}(t)(\Delta H)^{\mathrm{T}}]$ and Φ_1 is given by (21). Thus, if $\Phi_1 < 0$ in (21), we have

$$\dot{V}(t) \leq -2\delta V(t),\ t \geq 0, \tag{40}$$

which with $I_n \leq P \leq pI_n$ yields

$$|z(t)|^2 \leq V(t) \leq e^{-2\delta t}V(0) \leq pe^{-2\delta t}|z(0)|^2,$$

namely,

$$|z(t)| \leq \sqrt{p}e^{-\delta t}|z(0)|,\ t \geq 0. \tag{41}$$

On the other hand, by using the variation of constants formula for (38), we have

$$z(t) = e^{KHt}z(0),\ t \geq 0. \tag{42}$$

By norm's definition and (41)-(42), we obtain

$$\begin{aligned}\left\|e^{KHt}\right\| &= \sup_{|z(0)|=1}\left|e^{KHt}z(0)\right| \\ &\overset{(42)}{=} \sup_{|z(0)|=1}|z(t)| \overset{(41)}{\leq} \sqrt{p}e^{-\delta t}.\end{aligned} \tag{43}$$

With (43), inequality (37) can be continued as

$$\begin{aligned}|z(t)| &< \sqrt{p}e^{-\delta(t-\varepsilon)}|z(\varepsilon)| \\ &\quad + \varepsilon\Delta(\Delta_1 + \Delta_2 + \Delta_3)\sqrt{p}\textstyle\int_{\varepsilon}^{t} e^{-\delta(t-s)}ds \\ &= \sqrt{p}e^{-\delta(t-\varepsilon)}|z(\varepsilon)| \\ &\quad + \frac{\varepsilon\Delta(\Delta_1+\Delta_2+\Delta_3)\sqrt{p}}{\delta}\left(1 - e^{-\delta(t-\varepsilon)}\right) \\ &\leq \sqrt{p}e^{-\delta(t-\varepsilon)}|z(\varepsilon)| + \frac{\varepsilon\Delta(\Delta_1+\Delta_2+\Delta_3)\sqrt{p}}{\delta}.\end{aligned} \tag{44}$$

Note from (18), (31) and (33) that

$$\begin{aligned}|z(\varepsilon)| &= |\tilde{\theta}(\varepsilon) - G(\varepsilon)| \leq |\tilde{\theta}(\varepsilon)| + |G(\varepsilon)| \\ &< |\tilde{\theta}(0)| + \varepsilon\Delta + \tfrac{\varepsilon\Delta}{2} = |\tilde{\theta}(0)| + \tfrac{3\varepsilon\Delta}{2},\end{aligned}$$

by which, inequality (44) for $t \geq \varepsilon$ can be continued as

$$|z(t)| < \sqrt{p}e^{-\delta(t-\varepsilon)}\left(|\tilde{\theta}(0)| + \tfrac{3\varepsilon\Delta}{2}\right) + \tfrac{\varepsilon\Delta(\Delta_1+\Delta_2+\Delta_3)\sqrt{p}}{\delta}.$$

Then

$$\begin{aligned}\left|\tilde{\theta}(t)\right| &= |z(t) + G(t)| \leq |z(t)| + |G(t)| \\ &< \sqrt{p}e^{-\delta(t-\varepsilon)}\left(|\tilde{\theta}(0)| + \tfrac{3\varepsilon\Delta}{2}\right) \\ &\quad + \tfrac{\varepsilon\Delta(\Delta_1+\Delta_2+\Delta_3)\sqrt{p}}{\delta} + \tfrac{\varepsilon\Delta}{2},\ t \geq \varepsilon,\end{aligned}$$

which implies the second inequality in (23) due to

$$\sqrt{p}\left(\sigma_0 + \tfrac{\varepsilon^*\Delta[2(\Delta_1+\Delta_2+\Delta_3)+3\delta]}{2\delta}\right) + \tfrac{\varepsilon^*\Delta}{2} < \sigma,$$

namely,

$$\sqrt{p}\left(\sigma_0 + \tfrac{\varepsilon^*\Delta[2(\Delta_1+\Delta_2+\Delta_3)+3\delta]}{2\delta}\right) < \sigma - \tfrac{\varepsilon^*\Delta}{2}.$$

The latter, by squaring of both sides, is equivalent to $\Phi_2 < 0$ in (21).

By contradiction-based arguments in [19] (see Appendix A), it can be proved that (21) results in (30). The proof is finished.

References

[1] P. Binetti, K.B. Ariyur, M. Krstic, and F. Bernelli, "Formation flight optimization using extremum seeking feedback", *Journal of guidance, control, and dynamics*, vol. 26, no. 1, pp. 132-142, 2003.

[2] Y.B. Salamah and U. Ozguner, "Distributed extremum-seeking for wind farm power maximization using sliding mode control", *Energies*, vol. 14, no. 4, p. 828, 2021.

[3] M. Abdelgalil and H. Taha, "Recursive averaging with application to bio-inspired 3-D source seeking", *Energies*, vol. 6, pp. 2816-2821, 2022.

[4] M. Krstic and H.-H. Wang, "Stability of extremum seeking feedback for general nonlinear dynamic systems", *Automatica*, vol. 36, no. 4, pp. 595-601, 2000.

[5] A. Scheinker and M. Krstic, *Model-free stabilization by extremum seeking.* Springer, 2017.

[6] Y. Tan, D. Nesic, and I. Mareels, "On non-local stability properties of extremum seeking control", *Automatica*, vol. 42, no. 6, pp. 889-903, 2006.

[7] Y. Tan, D. Nesic, I. Mareels, and A. Astolfi, "On global extremum seeking in the presence of local extrema", *Automatica*, vol. 45, no. 1, pp. 245-251, 2009.

[8] S. Liu and M. Krstic, "Stochastic averaging in discrete time and its applications to extremum seeking", *IEEE Transactions on Automatic control*, vol. 55, no. 10, pp. 2235-2250, 2010.

[9] S. Liu and M. Krstic, "Stochastic averaging in discrete time and its applications to extremum seeking", *IEEE Transactions on Automatic control*, vol. 61, no. 1, pp. 90-102, 2016.

[10] H.-B. Durr, M.S. Stankovic, C. Ebenbauer and K.H. Johansson, "Lie bracket approximation of extremum seeking systems", *Automatica*, vol. 49, no. 6, pp. 1538-1552, 2013.

[11] C. Labar, C. Ebenbauer, and L. Marconi, "ISS-like properties in Lie-bracket approximations and application to extremum seeking", *Automatica*, vol. 136, p. 110041, 2022.

[12] V. Grushkovskaya, A. Zuyev, and C. Ebenbauer, "On a class of generating vector fields for the extremum seeking problem: Lie bracket approximation and stability properties", *Automatica*, vol. 94, pp. 151-160, 2018.

[13] W. Moase, C. Manzie, and M. Brear, "Newton-like extremum-seeking for the control of thermoacoustic instability", *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2094-2105, 2010.

[14] A. Ghaffari, M. Krstic, and D. Nesic, "Multivariable Newton-based extremum seeking", *Automatica*, vol. 48, no. 8, pp. 1759-1767. 2012.

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

[16] H.K. Khalil, *Nonlinear Systems*. Upper Saddle River NJ: Prentice Hall, 2002.

[17] E. Fridman, and J. Zhang, "Averaging of linear systems with almost periodic coefficients: A time-delay approach", *Automatica*, vol. 122, p. 109287, 2020.

[18] X. Yang, J. Zhang, and E. Fridman, "Periodic averaging of discrete-time systems: A time-delay approach", *IEEE Transactions on Automatic Control*, vol. 68, no. 7, pp. 4482-4489, 2023.

[19] Y. Zhu and E. Fridman, "Extremum seeking via a time-delay approach to averaging", *Automatica*, vol. 135, p. 109965, 2022.

[20] Y. Zhu, E. Fridman, and T. Oliveira, "Sampled-data extremum seeking with constant delay: a time-delay approach", *IEEE Transactions on Automatic Control*, vol. 68, no. 1, pp. 432-439, 2023.