Universidade do Estado do Rio de Janeiro

Centro de Tecnologia e Ciências

Faculdade de Engenharia

Elly d'Alcantara Fonseca Holness

A Simplified Gradient-Based Extremum Seeking for Multivariable Static Maps with Different Input Time Delays

Rio de Janeiro

2020

Elly d'Alcantara Fonseca Holness

A Simplified Gradient-Based Extremum Seeking for Multivariable Static Maps with Different Input Time Delays

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre em Ciências, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Sistemas Inteligentes e Automação.

Orientador: Prof. Dr. Tiago Roux de Oliveira

Rio de Janeiro

2020

CATALOGAÇÃO NA FONTE

UERJ / REDE SIRIUS / BIBLIOTECA CTC/B

H753 Holness, Elly d'Alcantara Fonseca.
A simplified gradient-based extremum seeking for multivariable static maps with different input time delays / Elly d'Alcantara Fonseca Holness. – 2020.
65f.

Orientador: Tiago Roux de Oliveira.
Dissertação (Mestrado) – Universidade do Estado do Rio de Janeiro, Faculdade de Engenharia.

1. Engenharia Eletrônica - Teses. 2. Controladores programáveis - Teses. 3. Sistemas de controle ajustável - Teses. 4. Otimização matemática - Teses. 5. Sistemas de controle por realimentação - Teses. I. Oliveira, Tiago Roux de. II. Universidade do Estado do Rio de Janeiro, Faculdade de Engenharia. III. Título.

CDU 681.513.63

Bibliotecária: Júlia Vieira – CRB7/6022

Assinatura _____ Data _____

Elly d'Alcantara Fonseca Holness

A Simplified Gradient-Based Extremum Seeking for Multivariable Static Maps with Different Input Time Delays

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre em Ciências, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Sistemas Inteligentes e Automação.

Aprovado em: 18 de Fevereiro de 2020

Banca Examinadora:

Prof. Tiago Roux de Oliveira, D.Sc. (Orientador)

Programa de Pós-Graduação em Engenharia Eletrônica - UERJ

Prof. Diego Campos Knupp, D.Sc.

Instituto Politécnico - UERJ

Prof. Fernando Cesar Lizarralde, D.Sc.

Programa de Engenharia Elétrica - UFRJ

Prof. Tito Luís Maia Santos, D.Eng.

Escola Politécnica - UFBA

Rio de Janeiro

2020

AGRADECIMENTOS

Ao mestre e orientador, Professor Tiago Roux, por toda confiança e apoio despendidos na orientação deste trabalho. Obrigada por sempre me incentivar, cobrar, e persistir, sem sua ajuda este trabalho não existiria.

Ao Professor Daisuke Tsubakino da Universidade de Nayoga, Japão, autor do relatório técnico que inspirou este trabalho.

Ao meu esposo Jamie Holness, pelo amor e incentivo à conclusão deste projeto. Ao meu filho Benjamin, que mesmo tão pequeno, teve que aceitar, entender e apoiar sua mamãe que nem sempre esteve disponível para brincar e cuidar.

Ao meus pais, Eliana, João e Peggy que juntos formaram a mais importante rede de apoio para que eu conseguisse concluir este trabalho.

Aos meus amigos do trabalho que torceram e me incentivaram nessa conquista.

RESUMO

HOLNESS, Elly d'Alcantara Fonseca *Controle Extremal Simplificado Baseada no Gradiente para Mapas Estáticos Multivariáveis com Diferentes Atrasos de Entrada.* 65 f. Dissertação (Mestrado em Engenharia Eletrônica) - Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro (UERJ), Rio de Janeiro, 2020.

Este estudo aborda a análise e projeto de controle extremal multivariável para mapas estáticos sujeitos a atrasos arbitrariamente longos. O método do gradiente é considerado. São tratados os sistemas de múltiplas entradas com atrasos diferentes em cada canal de entrada. No método alternativo, a compensação de fase dos sinais de excitação e a inclusão de feedback do preditor com uma estimativa da Hessiana baseada em perturbação (baseada na média) permitem obter resultados de convergência exponencial local para uma pequena vizinhança do ponto ótimo, mesmo na presença de atrasos. A análise de estabilidade é realizada sem o uso de transformação backstepping, que também elimina a complexidade do controlador. Em suma, assegura-se um esquema de implementação mais simples e análise direta sem invocar sucessivas transformações de backstepping. Um exemplo númerico ilustra o desempenho do controle extremal com compensação de atraso e sua simplicidade.

Palavras-chave: Controle adaptativo; Método gradiente de controle extremal; Otimização em tempo real; Realimentação por preditor; Sistemas com atraso.

ABSTRACT

HOLNESS, Elly d'Alcantara Fonseca *A Simplified Gradient-Based Extremum Seeking for Multivariable Static Maps with Different Input Time Delays.* 65 f. Dissertação (Mestrado em Engenharia Eletrônica) - Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro (UERJ), Rio de Janeiro, 2020.

This study addresses the design and analysis of multivariable Extremum Seeking for static maps subject to arbitrarily long time delays. Gradient-based method is considered. Multi-input systems with different time delays in each individual input channel are dealt with. In the alternative method the phase compensation of the dither signals and the inclusion of predictor feedback with a perturbation-based (averaging-based) estimate of the Hessian allow to obtain local exponential convergence results to a small neighborhood of the optimal point, even in the presence of delays. The stability analysis is carried without using backstepping transformation, which also eliminates the complexity of the controller. In a nutshell, a simpler implementation scheme and direct analysis without invoking successive backstepping transformation can be assured. A numerical example illustrates the performance of the proposed delay-compensated extremum seeking scheme and its simplicity.

Keywords: Adaptive control; Gradient extremum seeking; Real-time optimization; Predictor feedback; Delay systems.

LIST OF FIGURES

LIST OF ACRONYMS

ES	Extremum Seeking
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation

SUMMARY

INTRODUCTION

Extremum seeking (ES) is a method for real-time non-model based optimization. The adaptive control method is used for tuning parameters when there is a nonlinearity in the control problem with a local maximum or minimum. The most popular ES approach relies on a small periodic excitation, usually sinusoidal, to disturb the parameters being tuned [2–8]. This approach quantifies the effects of the parameters on the output of the nonlinear map, then uses that information to generate the search of the optimal values.

This work analyses Extremum Seeking with delays, starting from a review of different approaches for Extremum Seeking schemes with delays in the literature. The final goal is to present an alternative gradient-based extremum seeking control for multivariables static maps with different input time delays. Simulation results based on this alternative method are also presented.

History

In recent years, there have been a lot of advances in theory and applications of ES. This list includes the proof of local [2–4] or semi-global [7] stability properties of the search algorithm even in the presence of local extrema [8], its extension to the multivariable case [9] and advances in parameter convergence and performance improvement [5,6,10,11]. The book [12] also presents stochastic versions of the algorithm with filtered noise perturbation signals.

A lot of processes in different sectors of industry have to deal with delays that interfere in the stability and performance of the systems. These delays can be caused by properties intrinsic to the process, in the input of the control signal (input delays), or in the measurement of the controlled variables (output delays). Additionally, delays can be constant of variant in time, known or unknown.

Recently, the infinite-dimensional backstepping transformation [13] had shed light on new predictor feedback designs for delay compensation. Such a methodology has opened the possibility to construct explicit Lyapunov functionals for stability analysis or to quantify the performance and the robustness of the predictor based control scheme.

The motivation for this study is that there are applications in which post-processing

of the plant's measured output translates into a considerable delay in generating the control input to be applied to the plant.

The work in [14] proposes a solution to the problem of multivariable ES algorithms for output and/or equal input delay systems via predictor feedback, presenting two approaches to construct a predictor via perturbation-based estimates of the model. The first one is based on gradient optimization where they estimate the Hessian ([9]; [15]) for the purpose of implementing a predictor that compensates the delay. The second approach is based on the Newton optimization where they estimate the Hessian's inverse for the purpose of making the convergence rate independent of the unknown parameters of the map.

The publication in [16] extends the result from [14] to multiple and distinct input delays, for both gradient and Newton methods. In order to compensate multiple distinct delays in multivariable ES, they have developed an extension of the predictor feedback approach for the multi-input case with distinct delays through the introduction of a new successive-backstepping transformation [17]. The result is a very complex predictor design with stability analysis being carried out by using a novel successive backstepping transformation and averaging in infinite dimensions.

The publication [16] extends the result to multiple and distinct input delays, for both gradient and Newton methods, resulting in a very complex predictor design with stability analysis being carried out by using a novel successive backstepping transformation and averaging in infinite dimensions. On the other hand, [18] proposes new designs for multivariable extremum seeking for static maps with arbitrarily long time delays. Such a sequential predictor based approach eliminates the need for distributed terms. However, in sequential predictors larger delays or higher-dimension maps require smaller values of the dither excitation frequencies and therefore lead to slower convergence.

In contrast to [16], an alternative gradient-based extremum seeking scheme is presented in this dissertation that does not require backstepping transformation, which also leads to a much more simple control feedback structure. Indeed, this is the first result for predictor-feedback extremum seeking (with distributed terms) where the backstepping transformation is not employed. Moreover, the advantages of this metodology over the results in [18] seems to be that the delays in our approach are independent of the dither frequency and system's dimension, which means that we do not need to consider smaller

delays (or lower-order maps) to achieve faster convergence rates.

Methodology

The methodology used during the development of this study is:

- Statement of the control problem. A static quadratic map will be used for simplicity, and the general premise is that the delays are known.
- Controller proposal based predictor feedback with perturbation-based estimate of the Hessian;
- Stability analysis for the average system to conclude local exponential convergence;
- Simulation for $\mathbb{R}^2 \to \mathbb{R}$ system. The size of the system was chosen for computational simplicity, as the controller equations in a $\mathbb{R}^2 \to \mathbb{R}$ system is already very complex.

Notation and Norms

Considering a non linear generic systems $\dot{x} = f(t, x, \epsilon)$, where $x \in \mathbb{R}^n$, $f(x, t, \epsilon)$ is periodic in t period T, which means, $f(t + T, x, \epsilon) = f(t, x, \epsilon)$. Then, for $\epsilon > 0$ sufficiently small, it is possible to obtain the medium model given by $\dot{x}_{\text{av}} = f_{\text{av}}(x_{\text{av}})$, with $f_{\text{av}}(x) = \frac{1}{T}\int_0^T f(\tau, x_{\text{av}}, 0)d\tau$, where x_{av} denotes the medium version of the state $x(t)$ [19].

[19] defines a vector function $f(t, \epsilon) \in \mathbb{R}$ that is said to be of order $\mathcal{O}(\epsilon)$ inside the interval $[t_1, t_2]$ if there are positive constants k and ϵ^* such that $|f(t, \epsilon)| \leq K\epsilon \quad \forall \epsilon \in [t_1, t_2]$ and $\forall t \in [t_1, t_2]$. Sometimes it is possible to estimate k and ϵ^* and with that quantify $\mathcal{O}(\epsilon)$. Otherwise $\mathcal{O}(\epsilon)$ should be considered the order of magnitude for ϵ sufficiently small.

Averaging Theorem for FDE [20]

Consider the system with delay

$$\dot{x} = f\Big(\frac{t}{\epsilon}, x_t\Big), \quad \text{para } t > 0$$

where ϵ is a real parameter, $x_t(\Theta) = x(t+\Theta)$ for $-r \leq \Theta \leq 0$. Let Ω be a neighbourhood of 0 in $X = C([-r, 0]; \mathbb{R}^n)$, the supremum normed Banach space of continuous functions from $[-r, 0]$ to $\mathbb{R}^n$. Suppose $f : \mathbb{R} \times \Omega \to \mathbb{R}^n$ is continuous.

For $\varphi \in \Omega$ we assume that $f(t, \varphi)$ is almost periodic in t uniformly with respect to φ in compact subsets of Ω and f has a continuous Fréchet derivative $\partial(f, \varphi)/\partial\varphi x$ in φ on $\mathbb{R} \times \Omega$.

If $y = y_0 \in \Omega$ is a exponentially stable equilibrium to the averaged system $\dot{y}(t) = f_o(y_t)$, para $t > 0$, where

$$f_0(\varphi) = \lim_{T\to\infty} (1/T) \int_0^T f(s, \varphi)ds,$$

then, for some $\epsilon_0 > 0$ and $0 \leq \epsilon \leq \epsilon_0$, there exists a single periodic solution $t \longmapsto x^*(t, \epsilon)$ to the system with the properties of being continuous in t and ϵ, satisfying $|x^*(t, \epsilon) - y_0| \leq \mathcal{O}(\epsilon)$ para $t \in \mathbb{R}$, and there exists $\rho > 0$ such that, if $x(.; \rho)$ is a periodic solution to the system with $x(s) = \varphi$ and $|\varphi - y_0| < \rho$, then $|x(t) - x^*(t, \epsilon)| \leq Ce^{-\gamma(t-s)}$, for $C > 0$ and $\gamma > 0$.

Important Inequalities

In this work the inequalities of Cauchy-Schuarz are frequenty used, therefore they are defined here in the domain $[0, D]$.

Cauchy-Schwarz Inequality

$$\int_0^D u(x, t)\omega(x, t)dx \leq ||u(t)||\,||\omega(t)|| \tag{1}$$

Young's Inequality

$$ab \leq \frac{\gamma}{2}a^2 + \frac{1}{2\gamma}b^2, \quad \gamma > 0. \tag{2}$$

Structure of the Study

This work is divided in four chapters aiming to review the different approaches of multivariable Extremum Seeking with delays.

Chapter 1 is a review of the Extremum Seeking method for scalar nonlinear static maps.

Chapter 2 will review the gradient approach for multivariable Extremum Seeking with output and/or equal input delays presented in [14], whereas Chapter 3 will present the extended solution for different input delays proposed in [16]. In both cases stability analysis is carried out by Lyapunov functionals for predictor feedback via backstepping transformation.

Chapter 4 will present the alternative scheme proposed by [1] for different input delays, with stability analysis and simulation results for a $\mathbb{R}^2 \to \mathbb{R}$ static map.

1 EXTREMUM SEEKING REVIEW

Extremum seeking (ES) control is a real-time, model independent adaptive control technique for tuning parameters to optimize an unknown nonlinear map. The most popular extremum seeking approach relies on a small periodic excitation, usually sinusoidal, to disturb the parameters being tuned. This approach quantifies the effects of the parameters on the output of the nonlinear map, then uses that information to generate the search of the optimal parameter values.

Figure 1 shows the simplest perturbation-based Extremum Seeking scheme for a quadractic multivariable map $y = Q(\theta) = Q^* + \frac{1}{2}(\theta - \theta^*)^T H(\theta - \theta^*)$, where θ is the input vector $\theta = [\theta_1, \theta_2, \cdots, \theta_n]^T$ and Q^* is the extremum.

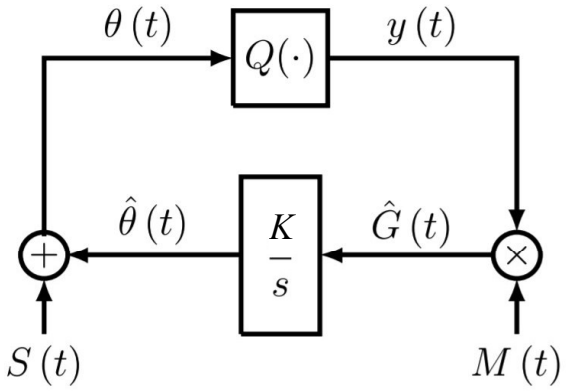

Figure 1 Simple perturbation-based Extremum Seeking scheme for an unknown multivariable map $y = Q(\cdot)$.

Three types of θ are presented in the control scheme, being θ^* the unknown optimizer of the map, $\hat{\theta}(t)$ the real time estimate of θ^* and $\theta(t)$ the actual input into the map. The optimal output y^* and Hessian H are also unknown. The user has to know whether the map has a maximum or a minimum i.e. the signal of the Hessian H.

The estimate $\hat{\theta}(t)$ is generated by the integrator K/s where the diagonal gain matrix K controls the speed of estimation for each input.

The dither signals that are used for the estimate the gradient are defined as

$$S(t) = \begin{bmatrix} a_1 \sin(\omega_1 t) & \ldots & a_n \sin(\omega_n t) \end{bmatrix}^T \tag{3}$$

$$M(t) = \left[\frac{2}{a_1} \sin(\omega_1 t) \quad ... \quad \frac{2}{a_n} \sin(\omega_n t) \right]^T \tag{4}$$

where $a_i \neq 0$. The sine wave was chosen for the perturbation signal, although a lot of other persistent perturbation could be used instead, from square waives to stochastic noise, as long as they have zero average. To guarantee convergence, $\omega_i \neq \omega_j$. And for simplicity in the convergence analysis, user should choose ω_i/ω_j as rational and $\omega_i + \omega_j \neq \omega_k$ for distinct i, j, and k.

The Extremum Seeking algorithm is successful if the error $\tilde{\theta}(t)$ between the estimate $\hat{\theta}(t)$ and the unknown θ^* converges towards zero.

$$\tilde{\theta}(t) = \hat{\theta}(t) - \theta^* \tag{5}$$

From Figure 1, $\dot{\hat{\theta}} = KM(t)y(t)$. Considering the quadratic map, the estimation error is governed by

$$\dot{\tilde{\theta}} = KM(t) \left[Q^* + \frac{1}{2}(\theta - \theta^*)^T H(\theta - \theta^*) \right] \tag{6}$$

But $\theta(t) = \hat{\theta}(t) + S(t)$ so by replacing right side of (6) in terms of $\tilde{\theta}$ in (5)we have

$$\dot{\tilde{\theta}} = KM(t)Q^* + \frac{KM(t)}{2}(\tilde{\theta} + S(t))^T H(\tilde{\theta} + S(t)) \tag{7}$$

Considering the sinusoidal components of $M(t)$ and $S(t)$, the averaged system is given by

$$\dot{\tilde{\theta}}_{\text{av}} = KH\tilde{\theta}_{\text{av}}. \tag{8}$$

If, for example, the map $Q(\cdot)$ has a maximum that is locally quadratic (which implies $H = H^T < 0$), and if user chooses the elements of the diagonal gain matrix K as positive, the gradient-based ES algorithm is guaranteed to be locally convergent. However, the convergence rate depends on the unknown Hessian H.

In the next chapters this basic scheme for Extremum Seeking Control will be extended to encompass delays in the control signal (input channels) or measurement of controlled variables (output channels), considering also multivariable problems.

1.1 Simulation Results

In order to evaluate the multidimentional version of the delay-free extremum seeking control, we consider the following static quadratic map, in the diagram of Figure 1 in a $n = 2$ example:

$$Q(\theta) = 1 + \frac{1}{2}\left(2(\theta_1)^2 + 4(\theta_2 - 1)^2 + 4\theta_1(\theta_2 - 1)\right) . \tag{9}$$

The extremum points are $\theta^* = (0, 1)$ and $y^* = 1$. Test was performed with diagonal $K = diag\{K_1, K_2\}$ where $K_1 = \frac{1}{100}$, $K_2 = \frac{1}{200}$, and parameters $a_1 = a_2 = 0.05$, $\omega = 0.5$, $\omega_1 = 17.5\omega$, $\omega_2 = 12.5\omega$, and $\hat{\theta}(0) = (1, 1)$.

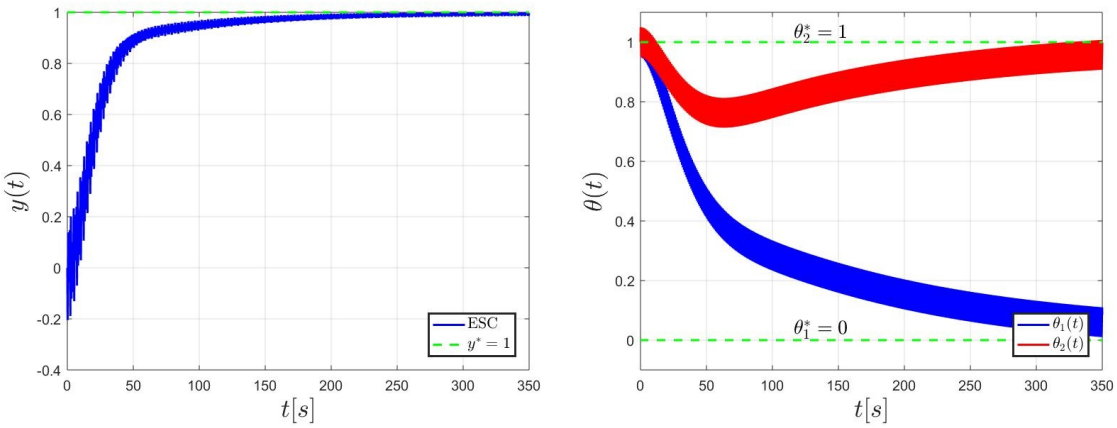

Figure 2 ES with no delays: time response for $y(t)$ and parameter $\theta(t)$

Figure 2 shows the time response for $y(t)$ and $\theta(t)$ respectively for the extremum seeking control with no delays. The control scheme properly stabilise the system.

2 MULTIVARIABLE GRADIENT ES WITH OUTPUT AND/OR EQUAL INPUT DELAYS

Multivariable Extremum Seeking control considers applications in which the goal is to maximize (or minimize) the scalar output $y \in \mathbb{R}$ of an unknown non-linear static map $y = Q(\theta)$ by varying the input vector $\theta = [\theta_1\, \theta_2\, \cdots\, \theta_n]^T$.

In this case we assume that a delay $D \geq 0$ is constant and known, and the measured output is given by

$$y(t) = Q(\theta(t - D)). \tag{10}$$

The system was assumed to be output delayed, however the results can be extended to the input-delay case, as long as they are the same in each individual input channel. The case when input delays D_{in} and output delays D_{out} occur simultaneously could also be handled, by assuming that the total delay to be counteract would be $D = D_{in} + D_{out}$ with D_{in}, $D_{out} \geq 0$.

Without loss of generality, let us consider the maximum seeking problem such that the maximizing value of θ is denoted by θ^*. For the sake of simplicity, we also assume that the non-linear map is quadratic, *i.e.*,

$$y(\theta) = y^* + \frac{1}{2}(\theta - \theta^*)^T H(\theta - \theta^*), \tag{11}$$

where besides the constants $\theta^* \in \mathbb{R}^n$ and $y^* \in \mathbb{R}$ being unknown, $H = H^T < 0$ is the $n \times n$ unknown Hessian matrix of the static map.

Let $\hat{\theta}$ be the estimate of θ^* and

$$\tilde{\theta}(t) = \hat{\theta}(t) - \theta^* \tag{12}$$

be the estimation error. From Figure 3, we obtain

$$G(t) = M(t)y(t) \tag{13}$$

$$\theta(t) = \hat{\theta}(t) + S(t) \tag{14}$$

where the vector dither signals are given by

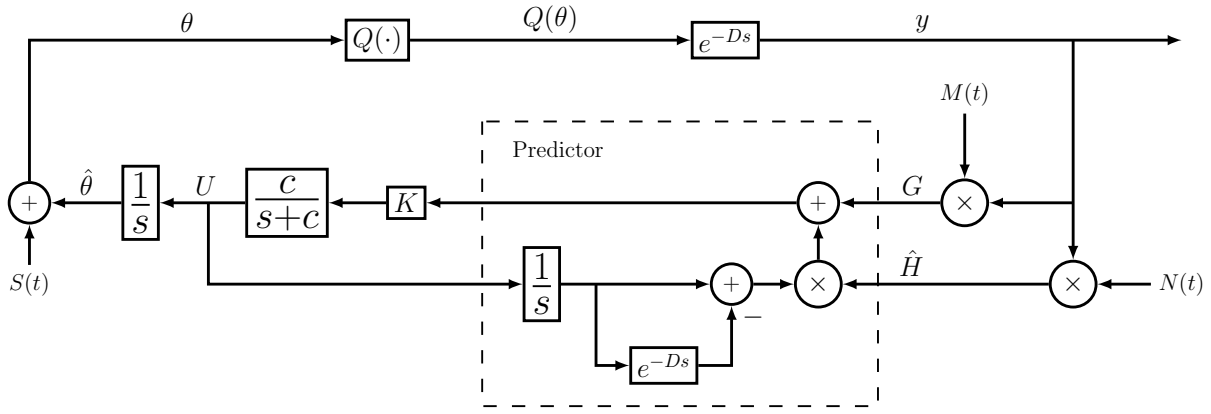

Figure 3 Block diagram of the basic prediction scheme for output-delay [1] compensation in multivariable gradient ES.

$$S(t) = \begin{bmatrix} a_1 \sin(\omega_1 t) & \dots & a_n \sin(\omega_n t) \end{bmatrix}^T \tag{15}$$

$$M(t) = \begin{bmatrix} \dfrac{2}{a_1} \sin(\omega_1 (t - D)) & \dots & \dfrac{2}{a_n} \sin(\omega_n (t - D)) \end{bmatrix}^T \tag{16}$$

with non-zero perturbation amplitudes a_i and frequencies $\omega_i \neq \omega_j$. For simplicity in the convergence analysis, we should choose ω_i/ω_j as a rational and $\omega_i + \omega_j \neq \omega_k$ for distinct i, j, k [9].

Notice that equation (16) is different from the multiplicative vector dither signal used in classical extremum seeking showed in Chapter 1. The phase shift $-\omega_i D$ is the first measure employed here to compensate the output delay in (10).

The probing frequencies ω_i's can be selected as

$$\omega_i = \omega_i' \omega = \mathcal{O}(\omega), \quad i \in 1, 2, \cdots, n, \tag{17}$$

where ω is a positive constant and ω_i' is a rational number. One possible choice is given in [9] as

$$\omega_i' \notin \left\{ \omega_j', \ \frac{1}{2}(\omega_j' + \omega_k'), \ \omega_j' + 2\omega_k', \ \omega_j' + \omega_k' \pm \omega_l' \right\}, \tag{18}$$

for all distinct i, j, k and l.

2.1 Averaging Analysis without Predictor Compensation

If the classical gradient Extremum Seeking control feedback law $U(t) = KG(t)$ was applied, one could write $\dot{\tilde{\theta}}(t) = \dot{\hat{\theta}}(t) = KG(t)$, where $K > 0$ is a $n \times n$ positive diagonal matrix. From (10), (13) and (14), the closed loop system would be written as:

$$\dot{\tilde{\theta}} = KM(t)Q(\theta^* + S(t) + \tilde{\theta}(t - D)), \tag{19}$$

in the parameter error variable (12). For the case of a quadratic map (11), by taking Π as

$$\Pi = 2\pi \times \text{LCM}\left\{\frac{1}{\omega_i}\right\}, \quad \forall i \in \{1, 2, \cdots, n\}, \tag{20}$$

where LCM stands for the least common multiple and using the following identities

$$\frac{1}{\Pi}\int_0^{\Pi} M(\sigma)Q^* d\sigma = 0, \tag{21}$$

$$\frac{1}{\Pi}\int_0^{\Pi} \frac{M(\sigma)}{2}(\tilde{\theta} + S(\sigma))^T H(\tilde{\theta} + S(\sigma))d\sigma = H\tilde{\theta}, \tag{22}$$

to average (19), we would obtain:

$$\frac{d\tilde{\theta}_{\text{av}}(t)}{dt} = KH\tilde{\theta}_{\text{av}}(t - D) \tag{23}$$

From (23), it is clear that the equilibrium $\tilde{\theta}^e_{\text{av}} = 0$ of the average system is not necessarily stable for arbitrary values of the delay D. This reinforces the necessity of applying the prediction $U(t) = KG(t + D)$, $\forall t \geq 0$ to stabilize the system.

2.2 Preditor Feedback via Hessian Estimation

The idea of the predictor feedback [14] is to compensate for the delay by feeding back the future state $G(t + D)$, or $G_{\text{av}}(t + D)$ in the equivalent average system. The average version of the vector signal (13) is given by

$$G_{\text{av}} = H\tilde{\theta}_{\text{av}}(t - D), \tag{24}$$

generated by the following average system

$$\dot{\tilde{\theta}}_{\text{av}}(t-D) = U_{\text{av}}(t-D), \tag{25}$$

since

$$\dot{\tilde{\theta}}(t-D) = U(t-D) \tag{26}$$

with $U \in \mathbb{R}^n$ and $U_{\text{av}} \in \mathbb{R}^n$. Consequently,

$$\dot{G}_{\text{av}}(t) = HU_{\text{av}}(t-D). \tag{27}$$

Given the stabilizing diagonal matrix $K > 0$ for the undelayed system, our wish is to have a controller that achieves

$$U_{\text{av}}(t) = KG_{\text{av}}(t+D), \quad \forall t \geq 0, \tag{28}$$

and it appears to be non implementable since it requires future values of the state. However, by applying the variation of constants formula to (27) we can express the future state as

$$G_{\text{av}}(t+D) = G_{\text{av}}(t) + H\int_{t}^{t+D} U_{\text{av}}(\sigma - D)d\sigma \tag{29}$$

where the current state $G_{\text{av}}(t)$ is the initial condition. Shifting the time variable under the integral in (29), we obtain

$$G_{\text{av}}(t+D) = G_{\text{av}}(t) + H\int_{t-D}^{t} U_{\text{av}}(\sigma)d\sigma, \tag{30}$$

which gives the future state $G_{\text{av}}(t+D)$ in terms of the average control signal $U_{\text{av}}(\sigma)$ from the past window $[t-D, t]$. It yields the following feedback law

$$U_{\text{av}}(t) = K\left[G_{\text{av}}(t) + H\int_{t-D}^{t} U_{\text{av}}(\sigma)d\sigma\right]. \tag{31}$$

Hence, from (30) and (31), the average feedback law (28) can be obtained indeed as desired. Consequently,

$$\dot{\tilde{\theta}}_{\text{av}}(t) = KG_{\text{av}}(t+D), \quad \forall t \geq 0, \tag{32}$$

Therefore, from (24), one has

$$\frac{d\tilde{\theta}_{\text{av}}(t)}{dt} = KH\tilde{\theta}_{\text{av}}(t), \quad \forall t \geq D, \tag{33}$$

with an exponentially attractive equilibrium $\tilde{\theta}^e_{\text{av}} = 0$, since $KH < 0$. It means that the delay is perfectly compensated in D seconds, namely, the system evolves as if the delay were absent after D seconds.

Next, we show that the control objectives can still be achieved if a simple modification of the above basic predictor-based controller, which employs a low-pass filter, is applied. In this case, we propose the following infinitedimensional and averaging-based predictor feedback in order to compensate the delay [21]

$$U(t) = \frac{c}{s+c}\left\{K\left[G(t) + \hat{H}(t)\int_{t-D}^{t} U(\tau)d\tau\right]\right\}, \quad K > 0 \tag{34}$$

where $c > 0$ is sufficiently large, *i.e.*, the predictor feedback is of the form of a low-pass filtered of the non average version of (31). This low pass filtering is particularly required in the stability analysis when the averaging theorem in infinite dimensions [20] is invoked. Basically, the output delay is handled as a state delay in the averaging theorem when we introduce the lag filter. Note that we mix the time and frequency domains in (34) by using the braces $\{\cdot\}$ to denote that the transfer function acts as an operator on a time-domain function.

The predictor feedback (34) is infinite-dimensional because the integral involves the control history over the interval $[t-D, t]$. On the other hand, it is averaging-base (perturbation-based) because $\hat{H}$ is updated according to the estimate of the unknown Hessian H:

$$\hat{H}(t) = N(t)y(t) \tag{35}$$

satisfying the following averaging property

$$\frac{1}{\Pi}\int_0^{\Pi} N(\sigma)yd\sigma = H, \tag{36}$$

demonstrated in [9] if a quadratic map as in (11) is considered. In other words, the average version $\hat{H}_{\text{av}} = (Ny)_{\text{av}} = H$.

The elements of the $n \times n$ demodulating matrix $N(t)$ for generating the estimate of the Hessian are given by:

$$N_{ij}(t) = \begin{cases} \frac{16}{a_i^2}\left(\sin^2(\omega_i(t-D)) - \frac{1}{2}\right), & i = j \\ \frac{4}{a_i a_j}\sin(\omega_i(t-D))\sin(\omega_j(t-D)), & i \neq j \end{cases} \tag{37}$$

2.3 Stability Analysis

The feedback law (34) seems to be implicit since U is present on both sides. However, the input memory $U(\sigma)$, where $\sigma \in [t-D, t]$, is part of the state of an infinite-dimensional system, and thus the control law is effectively a complete-state-feedback controller.

However, the analysis sketched previously does not capture the entire system consisting of the ODE in (27) and the infinite-dimensional subsystem of the input delay. The developments in [13] make it possible to address these concerns due to the availability of Lyapunov functions for predictor feedback via backstepping transformation. The backstepping construction permits a stability analysis of the complete feedback system with the cascade PDE-ODE described below and the infinite-dimensional control law (34), resulting in an exponential stability estimate in the appropriate norm of this system.

By applying averaging theorem in infinite dimensions [20], we can state the following stability result for gradient Extremum Seeking control in the presence of output delays.

Theorem 1. *Consider the control system in Figure 3 with delayed output (10) and non-linear map (11). There exists $c^* > 0$ such that, $\forall c \geq c^*$, $\exists \omega^*(c) > 0$ such that, $\forall \omega > \omega^*$, the closed-loop delayed system (26) and (34) with state $\tilde{\theta}(t-D)$, $U(\tau)$, $\forall \tau \in [t-D, t]$, has a unique exponentially stable periodic solution in t of period Π, denoted by $\tilde{\theta}^{\Pi}(t-D)$, $U^{\Pi}(\tau)$, $\forall \tau \in [t-D, t]$, satisfying, $\forall t \geq 0$:*

$$\left(\left|\tilde{\theta}^{\Pi}(t-D)\right|^2 + \left|U^{\Pi}(t)\right|^2 + \int_{t-D}^{t}\left|U^{\Pi}(\tau)\right|^2 d\tau\right)^{1/2} \leq \mathcal{O}(1/\omega). \tag{38}$$

Furthermore,

$$\limsup_{t\to+\infty} |\theta(t) - \theta^*| = \mathcal{O}(|a| + 1/\omega), \tag{39}$$

$$\limsup_{t\to+\infty} |y(t) - y^*| = \mathcal{O}(|a|^2 + 1/\omega^2), \tag{40}$$

where $a = [a_1\ a_3\ \cdots a_n]^T$.

Proof: The demonstration follows the **Steps 1** to **8** below.

Step 1: *Transport PDE for Delay Representation*

$$\dot{\tilde{\theta}}(t-D) = u(0,t), \tag{41}$$

$$u_t(x,t) = u_x(x,t), \quad x \in [0,D], \tag{42}$$

$$u(D,t) = U(t), \tag{43}$$

with solution

$$u(x,t) = U(t+x-D) \tag{44}$$

Step 2: *Equations of the Closed-loop System*

First, plug (12) and (14) into (11) so that the output is given in term of $\tilde{\theta}$:

$$y(t) = y^* + \frac{1}{2}(\tilde{\theta}(t-D) + S(t-D))^T H(\tilde{\theta}(t-D) + S(t-D)) \tag{45}$$

By plugging (13) and (35) into (34), and representing the integrand in (34) using the transport PDE state, one has

$$U(t) = \frac{c}{s+c}\left\{K\left[M(t)y(t) + N(t)y(t)\int_0^D u(\sigma,t)d\sigma\right]\right\}. \tag{46}$$

Finally, substituting (46) into (43), we can rewrite (41)-(43) as

$$\dot{\tilde{\theta}}(t-D) = u(0,t), \tag{47}$$

$$\partial_t u(x,t) = \partial_x u(x,t), \quad x \in [0,D], \tag{48}$$

$$u(D,t) = \frac{c}{s+c}\left\{K\left[M(t)y(t) + N(t)y(t)\int_0^D u(\sigma,t)d\sigma\right]\right\}, \tag{49}$$

with solution

$$u_{\text{av}}(x,t) = U_{av}(t+x-D). \tag{50}$$

Step 3: *Average Model of the Closed-loop System*

Analogously to the computation carried out in [9], the following two averaging properties can be obtained if a quadratic map as in (46) is considered:

$$\frac{1}{\Pi}\int_0^{\Pi} M(\lambda)yd\lambda = H\tilde{\theta}_{\text{av}}(t-D), \tag{51}$$

and

$$\frac{1}{\Pi}\int_0^{\Pi} N(\lambda)y\bar{u}d\lambda = H\int_0^{D} u_{\text{av}}(\sigma,t)d\sigma \tag{52}$$

where $\bar{u}(t) = \int_0^D u(\sigma,t)d\sigma$, whereas $\tilde{\theta}_{\text{av}}(t-D)$ and $u_{\text{av}}(\sigma,t)$ denote the average versions of $\tilde{\theta}(t-D)$ and $u(\sigma,t)$ respectively.

Now denoting

$$\tilde{\vartheta}(t) = \tilde{\theta}(t-D) \tag{53}$$

and using (34), the average version of system (47)-(49) is:

$$\dot{\tilde{\vartheta}}_{\text{av}}(t) = u_{\text{av}}(0,t), \tag{54}$$

$$\partial_t u_{\text{av}}(x,t) = \partial_x u_{\text{av}}(x,t), \quad x\in[0,D], \tag{55}$$

$$\frac{d}{dt}u_{\text{av}}(D,t) = -cu_{\text{av}}(D,t) + cKH\left[\tilde{\vartheta}_{\text{av}}(t) + \int_0^D u_{\text{av}}(\sigma,t)d\sigma\right], \tag{56}$$

where the filter $c/s+c$ is also in the state-space form.

Step 4: *Backstepping transformation, its inverse and the target system*

Now, consider the following backstepping transformation [13]

$$w(x,t) = u_{\text{av}}(x,t) - KH\left[\tilde{\vartheta}_{\text{av}}(t) + \int_0^x u_{\text{av}}(\sigma,t)d\sigma\right], \tag{57}$$

with which we want to map the system (54)-(56) into the target system:

$$\dot{\tilde{\vartheta}}_{\rm av}(t) = KH\tilde{\vartheta}_{\rm av}(t) + w(0,t), \tag{58}$$

$$w_t(x,t) = w_x(x,t), \quad x \in [0,D], \tag{59}$$

$$w(D,t) = -\frac{1}{c}\partial_t u_{\rm av}(D,t) \tag{60}$$

which is the cascade of the transport PDE subsystem (59) with boundary condition (60) and the exponentially stable system $\dot{\tilde{\vartheta}}_{\rm av} = KH\tilde{\vartheta}_{\rm av}$ in (58), reminding that $KH < 0$. Note that, using (57) for $x = D$ and the fact that $u_{\rm av}(D,t) = U_{\rm av}(t)$, from (60) we get (56), *i.e.*, the explicit equation of the average control version of (34)

$$U_{\rm av}(t) = \frac{c}{s+c}\left\{KH\left[\tilde{\vartheta}_{\rm av}(t) + \int_0^D u_{\rm av}(\sigma,t)d\sigma\right]\right\}. \tag{61}$$

It is easily seen that

$$w_t(D,t) = \partial_t u_{\rm av}(D,t) - KHu_{\rm av}(D,t), \tag{62}$$

where $\partial_t u_{\rm av}(d,t) = \dot{U}_{\rm av}(t)$. The inverse of (61) is given by

$$u_{\rm av}(x,t) = w(x,t) + KH\left[e^{KHx}\tilde{\vartheta}_{\rm av}(t) + \int_0^x e^{KH(x-\sigma)}w(\sigma,t)d\sigma\right]. \tag{63}$$

Applying (60) and (63) into (62) we get

$$w_t(D,t) = -cw(D,t) - KHw(D,t) - (KH)^2\left[e^{KHD}\tilde{\vartheta}_{\rm av}(t) + \int_0^D e^{KH(D-\sigma)}w(\sigma,t)d\sigma\right]. \tag{64}$$

Step 5: *Lyapunov-Krasovskii Funcional*

Since the transport PDE (59)-(60) is an exponentially stable system for $c > 0$ sufficiently large [21], the overall cascade (58)-(60) is exponentially stable. This fact is established with the Lyapunov functional [14]

$$V(t) = \tilde{\vartheta}_{\rm av}(t)^T P \tilde{\vartheta}_{\rm av}(t) + \frac{a}{2}\int_0^D (1+x) w(x,t)^T w(x,t) dx + \frac{1}{2} w(D,t)^T w(D,t), \qquad (65)$$

where $P = P^T > 0$ is the solution of the Lyapunov equation

$$P(KH) + (KH)^T P = -Q, \qquad (66)$$

for some $Q = Q^T > 0$, and the parameter $a > 0$ is to be chosen later. We have

$$\begin{aligned}
\dot{V}(t) &= \tilde{\vartheta}_{\rm av}^T ((KH)^T P + P(KH)) \tilde{\vartheta}_{\rm av} + 2\tilde{\vartheta}_{\rm av}^T P w(0,t) \\
&\quad + a\int_0^D (1+x) w(x,t)^T w_x(x,t) dx + w(D,t)^T w_t(D,t) \\
&= -\tilde{\vartheta}_{\rm av}^T Q \tilde{\vartheta}_{\rm av} + 2\tilde{\vartheta}_{\rm av}^T P w(0,t) + \frac{a(1+D)}{2} w(D,t)^T w(D,t) - \frac{a}{2} w(0,t)^T w(0,t) \\
&\quad - \frac{a}{2}\int_0^D w(x,t)^T w(x,t) dx + w(D,t)^T w_t(D,t) \\
&\le -\tilde{\vartheta}_{\rm av}^T Q \tilde{\vartheta}_{\rm av} + \frac{2}{a}\left|\tilde{\vartheta}_{\rm av}^T P\right|^2 - \frac{a}{2}\int_0^D w(x,t)^T w(x,t) dx \\
&\quad + w(D,t)^T \left[w_t(D,t) + \frac{a(1+D)}{2} w(D,t) \right].
\end{aligned}$$

Let us choose

$$a = 4\frac{\lambda_{\max}(P^2)}{\lambda_{\min}(Q)}. \qquad (67)$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are minimum and maximum eigenvalues of the corresponding matrices. Then,

$$\begin{aligned}
\dot{V}(t) &\le -\frac{1}{2}\lambda_{\min}(Q)\left|\tilde{\vartheta}_{\rm av}(t)\right|^2 - \frac{a}{2}\int_0^D w(x,t)^T w(x,t) dx \\
&\quad + w(D,t)^T \left[w_t(D,t) + \frac{a(1+D)}{2} w(D,t) \right]. \qquad (68)
\end{aligned}$$

Now we consider (68) along with (64). With a completion of squares, we obtain

$$\begin{aligned}\dot{V}(t) &\leq -\frac{1}{4}\lambda_{\min}(Q)\left|\tilde{\vartheta}_{\rm av}(t)\right|^2 - \frac{a}{4}\int_0^D w(x,t)^T w(x,t)dx \\ &+\frac{\left|(KH)^2 e^{KHD}\right|^2}{\lambda_{\min}(Q)}\left|w(D,t)\right|^2 + \frac{1}{a}\left\|(KH)^2 e^{KH(D-\sigma)}\right\|^2\left|w(D,t)\right|^2 \\ &+\left[\frac{a(1+D)}{2}+|KH|\right]\left|w(D,t)\right|^2 - c\left|w(D,t)\right|^2 . \end{aligned} \tag{69}$$

To obtain (69), we have used:

$$\begin{aligned}-w(D,t)^T\left\langle (KH)^2 e^{KH(D-\sigma)}, w(\sigma,t)\right\rangle &\leq \left|w(D,t)\right|\left\|(KH)^2 e^{KH(D-\sigma)}\right\|\|w(t)\| \\ &\leq \frac{a}{4}\|w(t)\|^2 + \frac{1}{a}\left\|(KH)^2 e^{KH(D-\sigma)}\right\|^2\left|w(D,t)\right|^2 ,\end{aligned} \tag{70}$$

where the first inequality is the Cauchy-Schwartz and the second is Young's, the notation $\langle\cdot\,,\cdot\rangle$ denotes the inner product in the spatial variable $\sigma\in[0\,,D]$, on which both $e^{KH(D-\sigma)}$ and $w(\sigma,t)$ depend, and $\|\cdot\|$ denotes the L_2 norm in σ, *i.e.*,

$$\begin{aligned}\left\|(KH)^2 e^{KH(D-\sigma)}\right\|^2 &= (KH)^2\int_0^D e^{2KH(D-\sigma)}d\sigma \\ \|w(t)\|^2 &= \int_0^D w(\sigma,t)^T w(\sigma,t)d\sigma = \int_0^D |w(\sigma,t)|^2 d\sigma\,.\end{aligned} \tag{71}$$

Then, from (69), we arrive at

$$\begin{aligned}\dot{V}(t) &\leq -\frac{1}{4}\lambda_{\min}(Q)\left|\tilde{\vartheta}_{\rm av}(t)\right|^2 - \frac{a}{4(1+D)}\int_0^D (1+x)w(x,t)^T w(x,t)dx \\ &-(c-c^*)w(D,t)^T w(D,t)\,,\end{aligned} \tag{72}$$

where

$$c^* = \frac{a(1+D)}{2}+|KH|+\frac{\left|(KH)^2 e^{KHD}\right|^2}{\lambda_{\min}(Q)}+\frac{1}{a}\left\|(KH)^2 e^{KH(D-\sigma)}\right\|^2 . \tag{73}$$

Hence, from (72), if c is chosen such that $c>c^*$, we obtain

$$\dot{V}(t) \leq -\mu V(t)\,, \tag{74}$$

for some $\mu > 0$. Thus, the closed-loop system is exponentially stable in the sense of the full state norm

$$\left(|\tilde{\vartheta}_{\rm av}(t)|^2 + \int_0^D w(x,t)^T w(x,t)dx + w(D,t)^T w(D,t)\right)^{1/2}, \tag{75}$$

i.e., in the transformed variable $(\tilde{\vartheta}_{\rm av}\,, w)$.

Step 6: *Exponential Stability Estimate (in L_2 norm) for the Average System (54)-(56)*

To obtain exponential stability in the sense of $\left(|\tilde{\vartheta}_{\rm av}(t)|^2 + \int_0^D |u_{\rm av}(x,t)|^2 dx + |u_{\rm av}(D,t)|^2\right)^{1/2}$ L_2 norm for the average system, we need to show that there exist positive numbers α_1 and α_2 such that

$$\alpha_1 \left(|\tilde{\vartheta}_{\rm av}(t)|^2 + \int_0^D |u_{\rm av}(x,t)|^2 dx + |u_{\rm av}(D,t)|^2\right) \leq V(t) \leq \alpha_2 \left(|\tilde{\vartheta}_{\rm av}(t)|^2 + \int_0^D |u_{\rm av}(x,t)|^2 dx + |u_{\rm av}(D,t)|^2\right). \tag{76}$$

This is straightforward to establish by using (57), (63), (65) and employing the Cauchy-Schwartz inequality and other calculations, as in the proof of Theorem 2.1 from [13].

Hence, with (63), we get

$$|\tilde{\vartheta}_{\rm av}(t)|^2 + \int_0^D |u_{\rm av}(x,t)|^2 dx + |u_{\rm av}(D,t)|^2 \leq \frac{\alpha_2}{\alpha_1} e^{-\mu t} \left(|\tilde{\vartheta}_{\rm av}(0)|^2 + \int_0^D |u_{\rm av}(x,0)|^2 dx + |u_{\rm av}(D,0)|^2\right), \tag{77}$$

which completes the proof of exponential stability.

Step 7: *Invoking Averaging Theorem*

To invoke the averaging theorem, first note that the closed-loop system (26) and

(34) can be rewritten as:

$$\dot{\tilde{\theta}}(t-D) = U(t-D)\,, \tag{78}$$

$$\dot{U}(t) = -cU(t) + c\left\{K\left[G(t) + \hat{H}(t)\int_{t-D}^{t} U(\tau)d\tau\right]\right\}\,, \tag{79}$$

where $z(t) = [\tilde{\theta}(t-D), U(t)]^T$ is the state vector. Moreover, from (13) and (35), one has

$$\dot{z}(t) = f(\omega t, z_t)\,, \tag{80}$$

where $z_t(\Theta) = z(t+\Theta)$ for $-D \leq \Theta \leq 0$ and f is an appropriate continuous functional, such that the averaging theorem by [20] and [22] can be directly applied considering $\omega = 1/\epsilon$.

Moreover, from (77), the origin of the average closed-loop system (54)–(56) with transport PDE for delay representation is exponentially stable since there exist positive constants α and β such that all solutions satisfy

$$\Psi(t) \leq \alpha e^{-\beta t}\Psi(0)\,, \quad \forall t \geq 0\,, \tag{81}$$

where $\Psi(t) \triangleq |\tilde{\vartheta}_{\rm av}(t)|^2 + \int_0^D |u_{\rm av}(x,t)|^2 dx + |u_{\rm av}(D,t)|^2$, or equivalently,

$$\Psi(t) \triangleq |\tilde{\theta}_{\rm av}(t-D)|^2 + \int_{t-D}^{t} |U_{\rm av}(\tau)|^2\, d\tau + |U_{\rm av}(t)|^2\,, \tag{82}$$

using (53) and (50).

Then, according to the averaging theorem [20], for ω sufficiently large, (47)–(49) has a unique exponentially stable periodic solution around its equilibrium (origin) satisfying (38).

Step 8: *Asymptotic Convergence to a Neighborhood of the Extremum* (θ^*, y^*)

By using the change of variables (53) and then integrating both sides of (41) within the interval $[t, \sigma + D]$:

$$\tilde{\vartheta}(\sigma + D) = \tilde{\vartheta}(t) + \int_t^{\sigma+D} u(0,s)ds\,. \tag{83}$$

From (44), we can rewrite (83) in terms of U, namely

$$\tilde{\vartheta}(\sigma + D) = \tilde{\vartheta}(t) + \int_{t-D}^{\sigma} U(\tau) d\tau \,. \tag{84}$$

Now, note that

$$\tilde{\theta}(\sigma) = \tilde{\vartheta}(\sigma + D) \,, \quad \forall \sigma \in [t - D, t] \,. \tag{85}$$

Hence,

$$\tilde{\theta}(\sigma) = \tilde{\theta}(t - D) + \int_{t-D}^{\sigma} U(\tau) d\tau \,, \quad \forall \sigma \in [t - D, t] \,. \tag{86}$$

By applying the supremum norm in both sides of (86), we have

$$\begin{aligned}
\sup_{t-D \le \sigma \le t} \left| \tilde{\theta}(\sigma) \right| &= \sup_{t-D \le \sigma \le t} \left| \tilde{\theta}(t - D) \right| + \sup_{t-D \le \sigma \le t} \left| \int_{t-D}^{\sigma} U(\tau) d\tau \right| \\
&\le \sup_{t-D \le \sigma \le t} \left| \tilde{\theta}(t - D) \right| + \sup_{t-D \le \sigma \le t} \int_{t-D}^{t} |U(\tau)| \, d\tau \\
&\le \left| \tilde{\theta}(t - D) \right| + \int_{t-D}^{t} |U(\tau)| \, d\tau \qquad \text{(by applying Cauchy-Schwarz inequality)} \\
&\le \left| \tilde{\theta}(t - D) \right| + \left(\int_{t-D}^{t} 1 \cdot d\tau \right)^{1/2} \times \left(\int_{t-D}^{t} |U(\tau)|^2 \, d\tau \right)^{1/2} \\
&\le \left| \tilde{\theta}(t - D) \right| + \sqrt{D} \left(\int_{t-D}^{t} |U(\tau)|^2 \, d\tau \right)^{1/2} .
\end{aligned} \tag{87}$$

Now, it is easy to check

$$\left| \tilde{\theta}(t - D) \right| \le \left(\left| \tilde{\theta}(t - D) \right|^2 + \int_{t-D}^{t} |U(\tau)|^2 \, d\tau \right)^{1/2} , \tag{88}$$

$$\left(\int_{t-D}^{t} |U(\tau)|^2 \, d\tau \right)^{1/2} \le \left(\left| \tilde{\theta}(t - D) \right|^2 + \int_{t-D}^{t} |U(\tau)|^2 \, d\tau \right)^{1/2} . \tag{89}$$

By using (88) and (89), one has

$$\left|\tilde{\theta}(t-D)\right| + \sqrt{D}\left(\int_{t-D}^{t} |U(\tau)|^2 \, d\tau\right)^{1/2} \leq (1+\sqrt{D})\left(\left|\tilde{\theta}(t-D)\right|^2 + \int_{t-D}^{t} |U(\tau)|^2 \, d\tau\right)^{1/2}. \tag{90}$$

From (87), it is straightforward to conclude that

$$\sup_{t-D\leq\sigma\leq t}\left|\tilde{\theta}(\sigma)\right| \;\leq\; (1+\sqrt{D})\left(\left|\tilde{\theta}(t-D)\right|^2 + \int_{t-D}^{t} |U(\tau)|^2 \, d\tau\right)^{1/2} \tag{91}$$

and, consequently,

$$\left|\tilde{\theta}(t)\right| \;\leq\; (1+\sqrt{D})\left(\left|\tilde{\theta}(t-D)\right|^2 + \int_{t-D}^{t} |U(\tau)|^2 \, d\tau\right)^{1/2}. \tag{92}$$

Inequality (92) can be given in terms of the periodic solution $\tilde{\theta}^{\Pi}(t-D)$, $U^{\Pi}(\sigma)$, $\forall\sigma \in [t-D, t]$ as follows

$$\begin{aligned}\left|\tilde{\theta}(t)\right| \;\leq\; & (1+\sqrt{D})\left(\left|\tilde{\theta}(t-D) - \tilde{\theta}^{\Pi}(t-D) + \tilde{\theta}^{\Pi}(t-D)\right|^2\right. \\ & \left. + \int_{t-D}^{t} \left|U(\tau) - U^{\Pi}(\tau) + U^{\Pi}(\tau)\right|^2 d\tau\right)^{1/2}.\end{aligned} \tag{93}$$

By applying Young's inequality and some algebra, the right-hand side of (93) and $\left|\tilde{\theta}(t)\right|$ can be majorized by

$$\begin{aligned}\left|\tilde{\theta}(t)\right| \;\leq\; & \sqrt{2}\,(1+\sqrt{D})\left(\left|\tilde{\theta}(t-D) - \tilde{\theta}^{\Pi}(t-D)\right|^2 + \left|\tilde{\theta}^{\Pi}(t-D)\right|^2\right. \\ & \left. + \int_{t-D}^{t} \left|U(\tau) - U^{\Pi}(\tau)\right|^2 d\tau + \int_{t-D}^{t} \left|U^{\Pi}(\tau)\right|^2 d\tau\right)^{1/2}.\end{aligned} \tag{94}$$

According to the averaging theorem [20] and [22], we can conclude that

$$\tilde{\theta}(t-D) - \tilde{\theta}^{\Pi}(t-D) \;\; \rightarrow \;\; 0\,, \tag{95}$$

$$\int_{t-D}^{t} \left|U(\tau) - U^{\Pi}(\tau)\right|^2 d\tau \;\; \rightarrow \;\; 0\,, \tag{96}$$

exponentially. Hence,

$$\limsup_{t\to+\infty} |\tilde{\theta}(t)| \;\; = \;\; \sqrt{2}\,(1+\sqrt{D})\left(\left|\tilde{\theta}^{\Pi}(t-D)\right|^2 + \int_{t-D}^{t} \left|U^{\Pi}(\tau)\right|^2 d\tau\right)^{1/2}. \tag{97}$$

From (38) and (97), we can write

$$\limsup_{t\to+\infty} |\tilde{\theta}(t)| \;\; = \;\; \mathcal{O}(1/\omega)\,. \tag{98}$$

From (12) and reminding that $\theta(t) = \hat{\theta}(t) + S(t)$ with $S(t) = [a_1 \sin(\omega_1 t) \;\cdots\; a_n \sin(\omega_n t)]^T$, one has that

$$\theta(t) - \theta^* \;\; = \;\; \tilde{\theta}(t) + S(t)\,. \tag{99}$$

Since the first term in the right-hand side of (99) is ultimately of order $\mathcal{O}(1/\omega)$ and the second term is of order $\mathcal{O}(|a|)$, then

$$\limsup_{t\to+\infty} |\theta(t) - \theta^*| = \mathcal{O}(|a| + 1/\omega)\,. \tag{100}$$

Finally, from (10),(11) and (100), we get (40).

Thus, the average closed-loop system consisting of the plant (48), (49) and (50) with the average control (61) is exponentially stable at the origin.

For the user, inequalities (39) and (40) guarantee that, if a_i in (15) are chosen small and ω in (17) is large, the input $\theta(t)$ converges to a small interval around the unknown θ^* and $y(t)$ converges to the vicinity of the optimal output y^*.

2.3.1 Simulation Results

In order to illustrate the result of proposed controller to compensate delays, simulation is presented using the static quadratic map (11) with the Hessian $H = \left[\begin{smallmatrix} -100 & -30 \\ -20 & -20 \end{smallmatrix}\right]$ and its inverse given by $H^{-1} = \left[\begin{smallmatrix} -0.0182 & 0.0273 \\ 0.0273 & -0.0909 \end{smallmatrix}\right]$. The input-output map is subject to an output delay $D = 5s$ according to (10). This value of delay induces instability in the closed loop system if no prediction action is adopted. The extremum point is $Q^* = 100$ and its maximizer is given by $\theta^* = [2 \quad 4]^T$. The parameters used for simulation are the following: $a = [0.1 \quad 0.1]^T$, $\omega = 0.2$ rad/s, $\omega_1 = 70\omega$, $\omega_2 = 50\omega$, $\omega_h = 0.08$ rad/s, $\omega_l = 0.1$ rad/s and $\hat{\theta}(0) = [2.5 \quad 5]^T$.

The update law (35) is applied to Hessian's inverse estimation and the gain matrix $K = 10^{-4}\,\mathrm{diag}\{0.25, 0.25\}$. The speed convergence is dictated by the Hessian H. Despite the slow response result showed in Figure 4, one can note that the deleterious delay could be counteract and the elements of the Hessian estimate $\hat{H}$ converge to the actual values of H, as depicted in Figure 5. Moreover, the extremum $Q^* = 100$ is ultimately achieved as well as its maximizer $\theta^* = [2 \quad 4]^T$

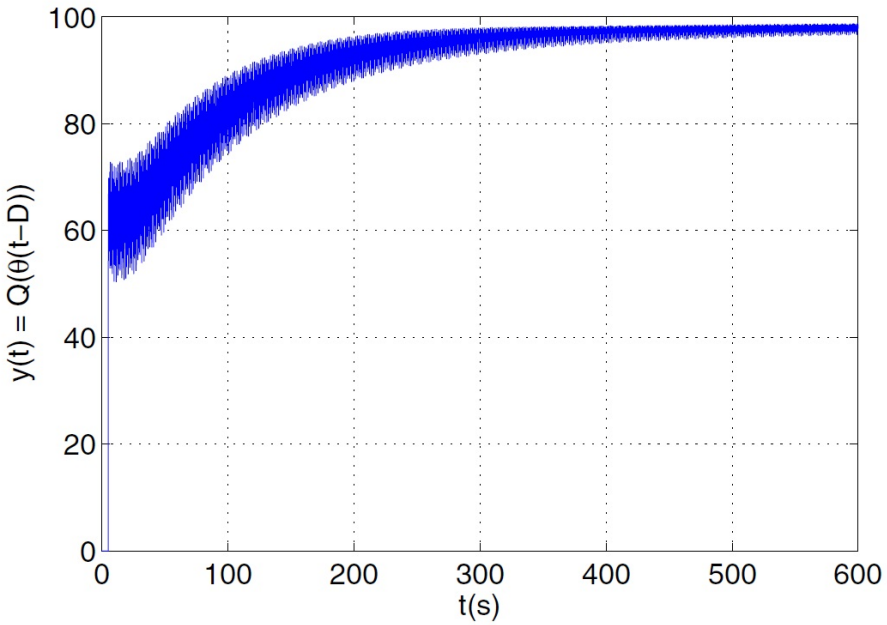

Figure 4 Gradient-based ES under output delay $D = 5s$: Time response of the delayed output $y(t)$ converging to the extremum $Q^* = 100$.

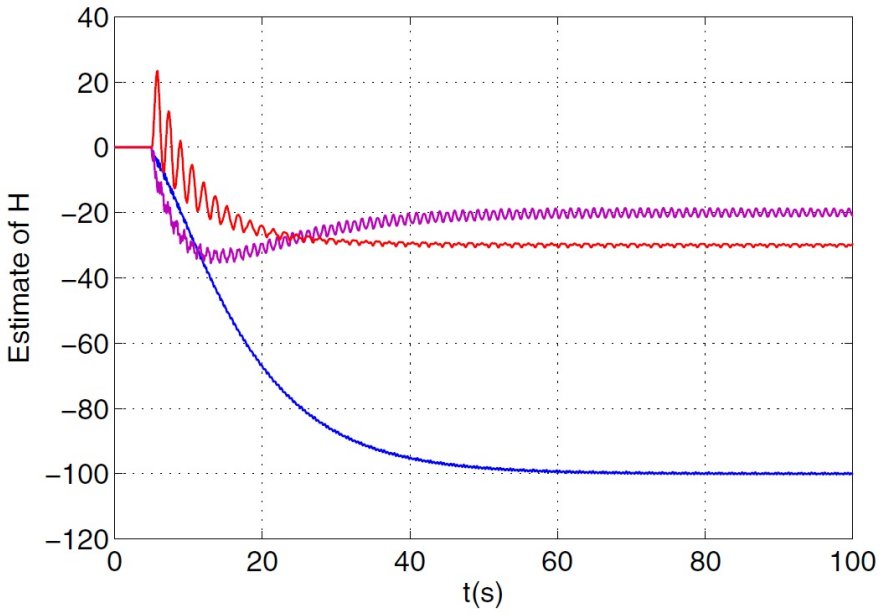

Figure 5 Time evolution of the elements of Hessian's estimator $\hat{H}(t)$ converging to the unknown values $(H)_{11} = -100$, $(H)_{12} = (H)_{21} = -20$, and $(H)_{22} = -30$.

3 GRADIENT-BASED ES WITH MULTIPLE AND DISTINCT INPUT DELAYS

In chapter 2 we reviewed the implementation of extremum seeking control for multivariable static maps subjected to output delays that, without loss of generality, can be extrapolated to multiple equal input delays.

In this chapter we will focus on multiple and distinct input delays, and the predictor from [16] will be presented.

We will assume the input-output representation (10) for the quadratic non-linear map (11). Thus, the locally quadratic static map with delay is given by

$$y(t) = y^* + \frac{1}{2}(\theta(t-D) - \theta^*)^T H(\theta(t-D) - \theta^*), \tag{101}$$

where the the delayed input vector can be represented by

$$\theta(t-D) := \begin{bmatrix} \theta_1(t-D_1) \\ \theta_2(t-D_2) \\ \vdots \\ \theta_n(t-D_n) \end{bmatrix}. \tag{102}$$

Without loss of generality we assume that the inputs have distinct delays which are ordered so that

$$D = diag\{D_1, \cdots, D_n\}, \quad 0 \leq D_1 \leq D_2 \leq \cdots \leq D_n. \tag{103}$$

In addition we consider that the constants D_i are known for all $i \in 1, 2, \cdots n$.

3.1 Gradient-Based ES with Multiple and Distinct Input Delays

From Figure 6 and $\tilde{\theta}(t)$ defined in (12), we can easily write

$$\dot{\tilde{\theta}}(t-D) = \begin{bmatrix} U_1(t-D_1) \\ U_2(t-D_2) \\ \vdots \\ U_n(t-D_n) \end{bmatrix}, \quad \dot{\tilde{\theta}}_i(t-D_i) = U_i(t-D_i)\,, \tag{104}$$

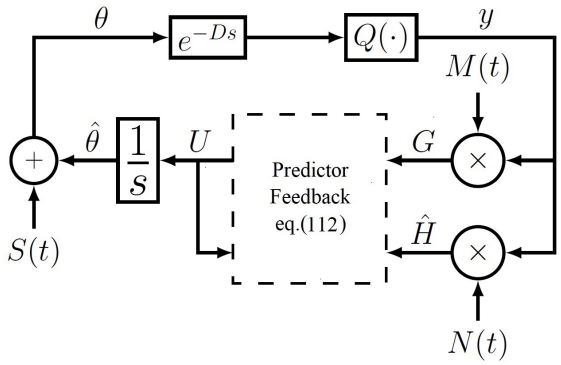

Figure 6 Block diagram of Gradient-based ES for multiple-input delay compensation

and the average model below by using (24):

$$\dot{G}_{\mathrm{av}}(t) = \sum_{i=1}^{n} H_i U_i^{\mathrm{av}}(t - D_i) = H \begin{bmatrix} U_1^{\mathrm{av}}(t - D_1) \\ U_2^{\mathrm{av}}(t - D_2) \\ \vdots \\ U_n^{\mathrm{av}}(t - D_n) \end{bmatrix}, \tag{105}$$

since

$$\dot{\hat{\theta}}_{\mathrm{av}}(t - D) = \begin{bmatrix} U_1^{\mathrm{av}}(t - D_1) \\ U_2^{\mathrm{av}}(t - D_2) \\ \vdots \\ U_n^{\mathrm{av}}(t - D_n) \end{bmatrix}. \tag{106}$$

In (105), $U_i^{\mathrm{av}}(t) \in \mathbb{R}$ for all $i \in \{1, 2, \ldots, n\}$ are the elements of the average control $U_{\mathrm{av}}(t) \in \mathbb{R}^n$ and the Hessian matrix $H = (H_1, H_2, \ldots, H_n) \in \mathbb{R}^{n \times n}$. In this case, there always exists a positive diagonal matrix $K = (K_1, K_2, \ldots, K_n)^T \in \mathbb{R}^{n \times n}$ such that HK is Hurwitz. For the sake of clarity, we say that H_i are column vectors of H and K_i are row vectors of K, $\forall i = 1, \ldots, n$.

Henceforth, the purpose of this section is to find a control feedback which has to perform prediction of the cross-coupling of the channels in (105). By applying the variation of constants formula to (105) gives

$$G_{\mathrm{av}}(t + D_i) = G_{\mathrm{av}}(t) + \sum_{j=1}^{n} \int_{t - D_j}^{t - (D_j - D_i)} H_j U_j^{\mathrm{av}}(\tau) d\tau, \tag{107}$$

such that the predictor feedback realizes

$$U_i^{\mathrm{av}}(t) = K_i^T G_{\mathrm{av}}(t + D_i)\,. \tag{108}$$

From (103), each control input U_i^{av} arrives at the plant at a different time. Since the system is causal, values of U_j^{av} with $j > i$ on the interval $(t - (D_j - D_i)\,, t)$ provide no information for prediction of $G_{\mathrm{av}}(t + D_i)$. On the other hand, values of U_j^{av} with $j < i$ on the interval $(t\,, t + D_i - D_j)$ are necessary to calculate (107), but they are unavailable future values. For $i\,, j \in \{1\,, \ldots\,, n\}$ with $i < j$, we have

$$\int_{t-D_j}^{t-(D_j-D_i)} (*)d\tau = \int_{t-D_j}^{t} (*)d\tau + \int_{t}^{t+D_i-D_j} (*)d\tau\,. \tag{109}$$

Differently from the results in Chapter 2, the variation of constants formula contains an extra term $\int_t^{t+D_i-D_j}(*)d\tau$.

In what follows, we will provide a predictor based controller that is more consistent with (107). To derive it, we have had to extend the backstepping approach [13] introducing a new successive backstepping-like transformation [17]. The explicit equation of this transformation of the delay state is given in the proof of the main theorem.

First of all, let us define the following notation:

$$A_i := \sum_{j=1}^{i} H_j K_j^T, \quad i \in \{0, 1, 2, \ldots, n\}\,, \tag{110}$$

being obvious that $A_0 = 0_{n\times n}$ and $A_n = HK$. In addition, the matrix-valued function Φ can be represented as [17]

$$\begin{aligned} \Phi(x,\zeta) &= e^{A_{i-1}(x-D_{i-1})} e^{A_{i-2}(D_{i-1}-D_{i-2})} \cdots e^{A_{j-1}(D_j-\zeta)}, \\ & D_{i-1} \le x < D_i, \qquad D_{j-1} \le \zeta < D_j, \end{aligned} \tag{111}$$

for any $i, j \in \{1, 2, \ldots, n\}$ satisfying $i > j$, and

$$\Phi(x,\zeta) = e^{A_{i-1}(x-\zeta)}, \qquad D_{i-1} \le \zeta \le x \le D_i, \tag{112}$$

$i \in \{1, 2, \ldots, n\}$, where we need to treat D_0 as 0.

In a few words, Φ can be seen as the state-transition matrix of the time varying system $\dot{X}(t) = A(t)X(t)$, $\forall t \geq 0$, where $A(t) \in \mathbb{R}^{n\times n}$ is a piecewise constant function defined by

$$A(t) = \begin{cases} 0_{n\times n}, & t \leq D_1, \\ A_i, & D_i < t \leq D_{i+1}, \quad i = 1, \ldots, n-1, \\ A_n, & t > D_n\,. \end{cases} \tag{113}$$

As it will be shown in the next section, the following predictor-based controller

$$\begin{aligned} U_i(t) = \frac{c}{s+c} \Bigg\{ & K_i^T \hat{\Phi}(D_i, 0) G(t) \\ & + K_i^T \left(\sum_{j=1}^{i} \int_{t-D_j}^{t} \hat{\Phi}(D_i, \tau - t + D_j) \hat{H}_j U_j(\tau) d\tau \right. \\ & \left. + \sum_{j=i+1}^{n} \int_{t-D_i}^{t} \hat{\Phi}(D_i, \tau - t + D_i) \hat{H}_j U_j(\tau - (D_j - D_i)) d\tau \right) \Bigg\}, \end{aligned} \tag{114}$$

guarantees exponential stability for the closed-loop system, with $\hat{\Phi}$ defined as in (111) and (112) but replacing A_i in (110) by $\hat{A}_i(t) := \sum_{j=1}^{i} \hat{H}_j(t) K_j^T$, vector $G(t)$ given in (13) and $\hat{H}$ being the columns of the Hessian estimate $\hat{H} = (\hat{H}_1, \hat{H}_2, \ldots, \hat{H}_n) \in \mathbb{R}^{n\times n}$ given by (35). For ω in (17) sufficiently large and a_i in (15)-(16) sufficiently small, the average version of the predictor-feedback form (114) can be numerically approximated by

$$\begin{aligned} U_i^{\text{av}}(t) = \frac{c}{s+c} \Bigg\{ & K_i^T \Phi(D_i, 0) G_{\text{av}}(t) \\ & + K_i^T \left(\sum_{j=1}^{i} \int_{t-D_j}^{t} \Phi(D_i, \tau - t + D_j) H_j U_j^{\text{av}}(\tau) d\tau \right. \\ & \left. + \sum_{j=i+1}^{n} \int_{t-D_i}^{t} \Phi(D_i, \tau - t + D_i) H_j U_j^{\text{av}}(\tau - (D_j - D_i)) d\tau \right) \Bigg\} \end{aligned} \tag{115}$$

From (113), it is possible to show the term between braces in (115) by itself, if applied to (105), is enough to conclude that the closed-loop system $\dot{G}_{\text{av}}(t) = A_n G_{\text{av}}(t)$ is totally delay-compensated $\forall t \geq D_n$, since $A_n = HK$ is Hurwitz. However, due to technical limitations involving averaging results in infinite dimensions, we must include a low-pass filter $c/(s+c)$ in the predictor control loop, as was done in (114) and (115).

3.2 Stability Analysis

The availability of Lyapunov functionals for predictor feedback via backstepping transformation [13] permits the stability analysis of the complete feedback system under delays with a cascade representation of ODE–PDE equations and the infinite-dimensional control law.

The resulting exponential stability estimate in L_2-norm of the closed-loop infinite-dimensional system is stated in the next theorem for Gradient ES subject to multiple and distinct input delays.

Theorem 2. *Consider the control system in Figure 6 with multiple and distinct input delays according to (10) and (104) and locally quadratic nonlinear map (11). There exists* $c^* > 0$ *such that,* $\forall c \geq c^*$, $\exists\ \omega^*(c) > 0$ *such that,* $\forall \omega > \omega^*$, *the closed-loop delayed system (104) and (114) with state* $\tilde{\theta}_i(t - D_i)$, $U_i(\tau)$, $\forall \tau \in [t - D_i, t]$ *and* $\forall i \in 1, 2, \ldots, n$, *has a unique locally exponentially stable periodic solution in* t *of period* Π, *denoted by* $\tilde{\theta}_i^{\Pi}(t - D_i)$, $U_i^{\Pi}(\tau)$, $\forall \tau \in [t - D_i, t]$ *and* $\forall i \in \{1, 2, \ldots, n\}$, *satisfying,* $\forall t \geq 0$*:*

$$\left(\sum_{i=1}^{n} \left[\tilde{\theta}_i^{\Pi}(t - D_i) \right]^2 + \left[U_i^{\Pi}(t) \right]^2 + \int_{t-D_i}^{t} \left[U_i^{\Pi}(\tau) \right]^2 d\tau \right)^{1/2} \leq \mathcal{O}(1/\omega)\,. \tag{116}$$

Furthermore,

$$\limsup_{t \to +\infty} |\theta(t) - \theta^*| = \mathcal{O}(|a| + 1/\omega)\,, \tag{117}$$

$$\limsup_{t \to +\infty} |y(t) - y^*| = \mathcal{O}(|a|^2 + 1/\omega^2)\,, \tag{118}$$

where $a = [a_1\ a_2\ \cdots\ a_n]^T$.

Proof: The demonstration follows the **Steps 1** to **7** below.

Step 1: *Transport PDE for Delay Representation*

Each individual delay D_i in equation (104) can be represented using a transport PDE as

$$\dot{\tilde{\theta}}_i(t - D_i) = u_i(0, t)\,, \tag{119}$$

$$\partial_t u_i(x, t) = \partial_x u_i(x, t)\,, \quad x \in [0, D_i]\,, \tag{120}$$

$$u_i(D_i, t) = U_i(t)\,, \quad i = 1, 2, \ldots, n\,, \tag{121}$$

where the solution of (120)–(121) is

$$u_i(x,t) = U_i(t + x - D_i) \tag{122}$$

and $u(x,t) = [u_1(x,t), \ldots, u_n(x,t)]^T$ is the state of the total delay infinite-dimensional subsystem.

Step 2: *Average Model of the Closed-loop System*

From (119)–(121), we can rewrite (105) as

$$\begin{aligned} \dot{G}_{\text{av}}(t) &= \sum_{i=1}^{n} H_i u_i^{\text{av}}(0,t), && (123)\\ \partial_t u_i^{\text{av}}(x,t) &= \partial_x u_i^{\text{av}}(x,t), \quad x \in [0, D_i], && (124)\\ u_i^{\text{av}}(D_i,t) &= U_i^{\text{av}}(t), \quad i = 1, 2, \ldots, n\,, && (125) \end{aligned}$$

where the solution of (124)–(125) is

$$u_i^{\text{av}}(x,t) = U_i^{\text{av}}(t + x - D_i) \tag{126}$$

and the PDE state is $u_{\text{av}}(x,t) = [u_1^{\text{av}}(x,t), \ldots, u_n^{\text{av}}(x,t)]^T$.

By representing the integrand in (115) using the transport PDE state, one has the average control law

$$\begin{aligned} U_i^{\text{av}}(t) = \frac{c}{s+c} \Bigg\{ K_i^T \Bigg(& \Phi(D_i, 0) G_{\text{av}}(t) \\ & + \sum_{j=1}^{n} \int_0^{\phi_j(D_i)} \Phi(D_i, \sigma) H_j u_j^{\text{av}}(\sigma, t) d\sigma \Bigg) \Bigg\}\,, \end{aligned} \tag{127}$$

with $\phi_j : [0, D_n] \to [0, D_j]$, $j \in \{1, 2, \ldots, n\}$ being the function defined by

$$\phi_j(x) = \begin{cases} x, & 0 \le x \le D_j, \\ D_j, & D_j < x < D_n\,. \end{cases} \tag{128}$$

Finally, substituting (127) into (125), we have

$$\dot{G}_{\rm av}(t) = \sum_{i=1}^{n} H_i u_i^{\rm av}(0,t), \tag{129}$$

$$\partial_t u_i^{\rm av}(x,t) = \partial_x u_i^{\rm av}(x,t), \quad x \in [0, D_i], \tag{130}$$

$$\begin{aligned}\frac{d}{dt}u_i^{\rm av}(D_i,t) = -c\ u_i^{\rm av}(D_i,t) + cK_i^T \Bigg(& \Phi(D_i,0)G_{\rm av}(t) \\ & + \sum_{j=1}^{n} \int_0^{\phi_j(D_i)} \Phi(D_i,\sigma) H_j u_j^{\rm av}(\sigma,t) d\sigma \Bigg) . \end{aligned} \tag{131}$$

Step 3: *Successive Backstepping-like transformation, its inverse and the target system*

Consider the new infinite-dimensional backstepping-like transformation [17] of the delay state

$$w_i(x,t) = u_i^{\rm av}(x,t) - K_i^T \left(\Phi(x,0)G_{\rm av}(t) + \sum_{j=1}^{n} \int_0^{\phi_j(x)} \Phi(x,\sigma) H_j u_j^{\rm av}(\sigma,t) d\sigma \right) \tag{132}$$

which maps the system (129)–(131) into the target system:

$$\dot{G}_{\rm av}(t) = A_n G_{\rm av}(t) + \sum_{i=1}^{n} H_i w_i(0,t)\,, \tag{133}$$

$$\partial_t w_i(x,t) = \partial_x w_i(x,t) - \sum_{j=1}^{i-1} \lambda_{ij}(x) w_j(D_j,t), \quad x \in [0, D_i]\,, \tag{134}$$

$$w_i(D_i,t) = -\frac{1}{c}\partial_t u_i^{\rm av}(D_i,t)\,, \quad i = 1,2,\ldots,n\,, \tag{135}$$

where $A_n = HK$ and the coefficients $\lambda_{ij} : [0, D_i] \to \mathbb{R}$ are

$$\lambda_{ij}(x) = \begin{cases} 0, & 0 \le x \le D_j, \\ K_i^T \Phi(x, D_j) H_j, & D_j < x \le D_i. \end{cases} \tag{136}$$

Note that the PDE for w_i is not a simple transport equation unless w_i vanishes at the right boundary. Using (132) for $x = D_i$ and the fact that $u_i^{\rm av}(D_i,t) = U_i^{\rm av}(t)$, we can directly obtain (131) and (127) from (135).

On the other hand, the inverse of (132) is given by

$$u_i^{\mathrm{av}}(x,t)=w_i(x,t)+K_i^T\left(e^{A_nx}G_{\mathrm{av}}(t)+\sum_{j=1}^{n}\int_0^{\phi_j(x)}e^{A_n(x-\sigma)}H_jw_j(\sigma,t)d\sigma\right). \tag{137}$$

For later use, now we find an expression for $\partial_t w_i(D_i,t)$. Differentiating (137) with respect to x on the interval $x\in(D_{i-1},D_i)$, gives

$$\begin{aligned}\partial_x u_i^{\mathrm{av}}(x,t) &= \partial_x w_i(x,t)+\sum_{j=i}^{n}K_i^TH_jw_j(x,t)\\ &\quad+K_i^TA_n\left(e^{A_nx}G_{\mathrm{av}}(t)+\sum_{j=1}^{n}\int_0^{\phi_j(x)}e^{A_n(x-\sigma)}H_jw_j(\sigma,t)dy\right).\end{aligned} \tag{138}$$

In light of (130)–(131) and (134)–(136), we arrive at

$$\begin{aligned}\partial_t w_i(D_i,t) &= -cw_i(D_i,t)-\sum_{j=1}^{i}K_i^T\Phi(D_i,D_j)H_jw_j(D_j,t)\\ &\quad-\sum_{j=i+1}^{n}K_i^TH_jw_j(D_i,t)-\gamma_i(0)^TG_{\mathrm{av}}(t)\\ &\quad-\sum_{j=1}^{n}\int_0^{\phi_j(D_i)}\gamma_i(\sigma)^TH_jw_j(\sigma,t)d\sigma,\end{aligned} \tag{139}$$

where $\gamma_i(x):=e^{A_n^T(D_i-x)}A_n^TK_i$ for each $i\in\{1,2,\dots,n\}$.

Note that the right-hand side contains $w_j(D_i,t)$ for each j greater than i, which is not a boundary value of w_j. For this reason, a key feature of the Lyapunov functional is the necessity of breaking the domain of integration for the terms $(1+x)w_i(x,t)^2$, as shown in the next step.

Step 4: *Lyapunov-Krasovskii Functional*

Let V be the candidate of Lyapunov function defined by

$$V(t)=G_{\mathrm{av}}(t)^TPG_{\mathrm{av}}(t)+\sum_{i=1}^{n}\sum_{j=1}^{i}\frac{\bar{a}_j}{2}\int_{D_{j-1}}^{D_j}(1+x)w_i(x,t)^2dx+\frac{1}{2}\sum_{i=1}^{n}w_i(D_i,t)^2, \tag{140}$$

where $P=P^T\in\mathbb{R}^{n\times n}$ is the solution of the Lyapunov equation $PA_n+A_n^TP=-Q$ for some $Q=Q^T>0$. The real constant $\bar{a}_1>0$ is determined later. The other constants $\bar{a}_2,\dots,\bar{a}_n$ are arbitrary real numbers satisfying $\bar{a}_1<\bar{a}_2<\cdots<\bar{a}_n$. To shorten notation,

we define a function $w : [0,1] \times [0,+\infty) \to \mathbb{R}^n$ by

$$w(\xi,t) = \begin{pmatrix} w_1(D_1\xi,t) & w_2(D_2\xi,t) & \cdots & w_n(D_n\xi,t) \end{pmatrix}^T, \tag{141}$$

for $0 \leq \xi \leq 1$. In addition, we omit the dependence on the temporal variable t for simplicity. For instance, we write V and $w_i(x)$ instead of $V(t)$ and $w_i(x,t)$. Differentiating V with respect to t yields

$$\begin{aligned}
\dot{V} &= -G_{\mathrm{av}}^T Q G_{\mathrm{av}} + 2G_{\mathrm{av}}^T PHw(0) \\
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{i} \frac{\bar{a}_j}{2} \int_{D_{j-1}}^{D_j} (1+x) 2w_i(x)\partial_t w_i(x) dx + w(1)^T \partial_t w(1) \\
&= -G_{\mathrm{av}}^T Q G_{\mathrm{av}} + 2G_{\mathrm{av}}^T PHw(0) - \frac{\bar{a}_1}{2} w(0)^T w(0) \\
&\quad - \sum_{i=2}^{n}\sum_{j=1}^{i-1} \frac{\alpha_j}{2} w_i(D_j)^2 + w(1)^T \partial_t w(1) \\
&\quad - \sum_{i=1}^{n}\sum_{\ell=1}^{i-1}\sum_{j=\ell+1}^{i} \bar{a}_j w_\ell(D_\ell) \int_{D_{j-1}}^{D_j} (1+x) K_i^T \Phi(x, D_\ell) H_\ell w_i(x) dx \\
&\quad + \frac{1}{2} w(1)^T \Delta w(1) - \sum_{i=1}^{n}\sum_{j=i}^{n} \frac{\bar{a}_i}{2} \int_{D_{i-1}}^{D_i} w_j(x)^2 dx,
\end{aligned} \tag{142}$$

where $\alpha_j > 0$ and $\Delta \in \mathbb{R}^{n\times n}$ are defined by

$$\alpha_j = (\bar{a}_{j+1} - \bar{a}_j)(1+D_j), \quad j \in \{1,2,\ldots,n-1\}, \tag{143}$$

$$\Delta = \mathrm{diag}\left\{\bar{a}_1(1+D_1), \bar{a}_2(1+D_2), \ldots, \bar{a}_n(1+D_n)\right\}. \tag{144}$$

In what follows we estimate the terms in each line of (142).

For the terms in the first line, we have

$$-G_{\mathrm{av}} Q G_{\mathrm{av}} + 2G_{\mathrm{av}}^T PHw(0) - \frac{\bar{a}_1}{2} w(0)^T w(0) \leq -G_{\mathrm{av}} \left(Q - \frac{2}{\bar{a}_1} PHH^T P \right) G_{\mathrm{av}}. \tag{145}$$

Setting $\bar{a}_1 = 4\lambda_{\max}(PHH^T P)/\lambda_{\min}(Q)$ leads to

$$-G_{\mathrm{av}} Q G_{\mathrm{av}} + 2G_{\mathrm{av}}^T PHw(0) - \frac{\bar{a}_1}{2} w(0)^T w(0) \leq -\frac{1}{2} G_{\mathrm{av}}^T Q G_{\mathrm{av}}. \tag{146}$$

Considering the second line of (142). It follows from (139) that

$$\begin{aligned}
&-\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{\alpha_j}{2}w_i(D_j)^2 + w(1)^T\partial_t w(1)\\
&= -cw(1)^Tw(1) - w(1)^TLw(1)\\
&\quad -\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{\alpha_j}{2}\Big(w_i(D_j)^2 + K_j^TH_iw_j(D_j)w_i(D_j)\Big)\\
&\quad - w(1)^T\Gamma(0)^TG_{\mathrm{av}} - \sum_{i=1}^{n}w_i(D_i)\sum_{j=1}^{n}\int_0^{\phi_j(D_i)}\gamma_i(\sigma)^TH_jw_j(\sigma)d\sigma, \qquad (147)
\end{aligned}$$

where $L=(L_{ij})$ is the $n\times n$ lower triangular matrix whose (i,j)th entry L_{ij} is given by

$$L_{ij} = \begin{cases} 0, & i<j,\\ K_i^T\Phi(D_i,D_j)H_j. & i\geq j. \end{cases} \qquad (148)$$

The matrix-valued function $\Gamma:[0,D_n]\to\mathbb{R}^{n\times n}$ is defined to be $\Gamma(\sigma)=(\gamma_1(\sigma),\gamma_2(\sigma),\ldots,\gamma_n(\sigma))$. By completing the square, we see that

$$\begin{aligned}
(147) \leq\ & -cw(1)^Tw(1) + \frac{1}{4}G_{\mathrm{av}}^TQG_{\mathrm{av}}\\
&-\frac{1}{2}w(1)\left(L+L^T+B-2\Gamma(0)^TQ^{-1}\Gamma(0)\right)w(1)\\
&-\sum_{i=1}^{n}w_i(D_i)\sum_{j=1}^{n}\int_0^{\phi_j(D_i)}\gamma_i(\sigma)^TH_jw_j(\sigma)d\sigma, \qquad (149)
\end{aligned}$$

where $B\in\mathbb{R}^{n\times n}$ is the diagonal matrix whose ith diagonal entry is

$$\frac{1}{\alpha_i}K_i^T\left(\sum_{j=i+1}^{n}H_jH_j^T\right)K_i, \quad i=1,2,\ldots,n. \qquad (150)$$

Substituting (146) and (149) into (142) leads to

$$\begin{aligned}\dot{V} \leq & -\frac{1}{4}G_{\rm av}^T Q G_{\rm av} - \frac{\bar{a}_1}{2}\sum_{i=1}^{n}\int_0^{D_i} w_i(x)^2 dx - cw(1)^T w(1) \\ & -\frac{1}{2}w(1)^T\left(L + L^T + B + \Delta - 2\Gamma(0)^T Q^{-1}\Gamma(0)\right)w(1) \\ & -\sum_{i=1}^{n} w_i(D_i)\sum_{j=1}^{n}\int_0^{\phi_j(D_i)} \gamma_i(\sigma)^T H_j w_j(\sigma) d\sigma \\ & -\sum_{i=1}^{n}\sum_{\ell=1}^{i-1}\sum_{j=\ell+1}^{i}\bar{a}_j w_\ell(D_\ell)\int_{D_{j-1}}^{D_j}(1+x)K_i^T\Phi(x,D_\ell)H_\ell w_i(x)dx. \end{aligned} \tag{151}$$

By using the Cauchy-Schwarz and Young's inequalities, we can show there exists a diagonal matrix $\Lambda \in \mathbb{R}^{n\times n}$ such that

$$\begin{aligned} & -\sum_{i=1}^{n} w_i(D_i)\sum_{j=1}^{n}\int_0^{\phi_j(D_i)} \gamma_i(\sigma)^T H_j w_j(\sigma) d\sigma \\ & -\sum_{i=1}^{n}\sum_{\ell=1}^{i-1}\sum_{j=\ell+1}^{i}\bar{a}_j w_\ell(D_\ell)\int_{D_{j-1}}^{D_j}(1+x)K_i^T\Phi(x,D_\ell)H_\ell w_i(x)dx \\ & \leq \frac{1}{2}w(1)^T\Lambda w(1) + \frac{\bar{a}_1}{4}\sum_{i=1}^{n}\int_0^{D_i} w_i(x)^2 dx. \end{aligned} \tag{152}$$

Then, we have

$$\dot{V} \leq -\frac{1}{4}G_{\rm av}^T Q G_{\rm av} - \frac{\bar{a}_1}{4}\sum_{i=1}^{n}\int_0^{D_i} w_i(x)^2 dx - w(1)^T(cI_{n\times n}+R)w(1), \tag{153}$$

where $R := \frac{1}{2}\left(L + L^T + B + \Delta - 2\Gamma(0)^T Q^{-1}\Gamma(0) + \Lambda\right)$. Hence, if $c > \lambda_{\min}(R)$, there exists $\mu > 0$ such that

$$\dot{V} \leq -\mu V\,. \tag{154}$$

Thus, the closed-loop system is exponentially stable in the sense of the full state norm

$$\left(G_{\rm av}(t)^T G_{\rm av}(t) + \sum_{i=1}^{n}\int_0^{D_i} w_i(x,t)^2 dx + w_i(D_i,t)^2\right)^{1/2}. \tag{155}$$

i.e., in the transformed variable $(G_{\rm av}\,, w)$.

Step 5: *Exponential Stability Estimate (in L_2 norm) for the Average System (129)–(131)*

To obtain exponential stability in the sense of the norm

$$\Upsilon(t) \triangleq \left(|G_{\rm av}(t)|^2 + \sum_{i=1}^{n} \int_0^{D_i} [u_i^{\rm av}(x,t)]^2 dx + [u_i^{\rm av}(D_i,t)]^2 \right)^{1/2}, \tag{156}$$

we must show there exist positive α_1 and α_2 such that

$$\alpha_1 \Upsilon(t)^2 \leq V(t) \leq \alpha_2 \Upsilon(t)^2 . \tag{157}$$

This is straightforward to establish by using (132), (137), (140) and employing the Cauchy-Schwartz inequality and other calculations, as in the proof of [13, Theorem 2.1].

Hence, with (154), we get

$$\begin{aligned} &|G_{\rm av}(t)|^2 + \sum_{i=1}^{n} \int_0^{D_i} [u_i^{\rm av}(x,t)]^2 dx + [u_i^{\rm av}(D_i,t)]^2 \\ &\leq \frac{\alpha_2}{\alpha_1} e^{-\mu t} \left(|G_{\rm av}(0)|^2 + \sum_{i=1}^{n} \int_0^{D_i} [u_i^{\rm av}(x,0)]^2 dx + [u_i^{\rm av}(D_i,0)]^2 \right), \end{aligned} \tag{158}$$

which completes the proof of exponential stability in the original variable $(G_{\rm av}\,, u_{\rm av})$.

Step 6: *Invoking Averaging Theorem*

First, note that the closed-loop system (104) and (114) can be rewritten as:

$$\dot{\tilde{\theta}}_i(t - D_i) = U_i(t - D_i)\,, \quad i = 1, \ldots, n\,, \tag{159}$$

$$\begin{aligned} U_i(t) = -cU_i(t) + cK_i^T \Bigg\{ & \hat{\Phi}(D_i, 0) G(t) \\ & + \sum_{j=1}^{i} \int_{t-D_j}^{t} \hat{\Phi}(D_i, \tau - t + D_j) \hat{H}_j U_j(\tau) d\tau \\ & + \sum_{j=i+1}^{n} \int_{t-D_i}^{t} \hat{\Phi}(D_i, \tau - t + D_i) \hat{H}_j U_j(\tau - (D_j - D_i)) d\tau \Bigg\} \end{aligned} \tag{160}$$

where $\eta(t) = [\tilde{\theta}(t - D), U(t)]^T$ is the state vector. Moreover, from the definitions of $\hat{\Phi}$ in (114) $G(t)$ in (13) and $\hat{H}(t)$ in (35), one can conclude they are implicit funcions of ωt such that (159) and (160) can be given in the next compact form

$$\dot{\eta}(t) \;\; = \;\; f(\omega t, \eta_t)\,, \tag{161}$$

where $\eta_t(\Theta) = \eta(t + \Theta)$ for $-D_n \leq \Theta \leq 0$ and f is an appropriate continuous functional,

such that the averaging theorem by [20,22] in Appendix can be directly applied considering $\omega = 1/\epsilon$.

From (158), the origin of the average closed-loop system (129)–(131) with transport PDE for delay representation is locally exponentially stable. Then, from (24) and (25), we can conclude the same results in the norm

$$\left(\sum_{i=1}^{n}\left[\tilde{\theta}_i^{\mathrm{av}}(t-D_i)\right]^2+\int_0^{D_i}[u_i^{\mathrm{av}}(x,t)]^2dx+[u_i^{\mathrm{av}}(D_i,t)]^2\right)^{1/2}$$

since H is non-singular, *i.e.*, $\left|\tilde{\theta}_i^{\mathrm{av}}(t-D_i)\right| \leq |H^{-1}|\,|G_{\mathrm{av}}(t)|$.

Thus, there exist positive constants α and β such that all solutions satisfy $\Psi(t) \leq \alpha e^{-\beta t}\Psi(0)$, $\forall t \geq 0$, where $\Psi(t) \triangleq \sum_{i=1}^{n}\left[\tilde{\theta}_i^{\mathrm{av}}(t-D_i)\right]^2 + \int_0^{D_i}[u_i^{\mathrm{av}}(x,t)]^2\,dx + [u_i^{\mathrm{av}}(D_i,t)]^2$, or equivalently,

$$\Psi(t) \triangleq \sum_{i=1}^{n}\left[\tilde{\theta}_i^{\mathrm{av}}(t-D_i)\right]^2 + \int_{t-D_i}^{t}[U_i^{\mathrm{av}}(\tau)]^2\,d\tau + [U_i^{\mathrm{av}}(t)]^2\,, \tag{162}$$

using (126). Then, according to the averaging theorem by [20, 22] in Appendix, for ω sufficiently large, (119)–(121) or (104) and (114), has a unique locally exponentially stable periodic solution around its equilibrium (origin) satisfying (116).

Step 7: *Asymptotic Convergence to the Extremum* (θ^*, y^*)

By using the change of variables $\tilde{\vartheta}_i(t) := \tilde{\theta}_i(t-D_i)$ and then integrating both sides of (119) within $[t, \sigma + D_i]$, we have:

$$\tilde{\vartheta}_i(\sigma+D_i) = \tilde{\vartheta}_i(t) + \int_t^{\sigma+D_i} u_i(0,s)ds\,, \qquad i=1,\ldots,n\,. \tag{163}$$

From (122), we can rewrite (163) in terms of U, namely

$$\tilde{\vartheta}_i(\sigma+D_i) = \tilde{\vartheta}_i(t) + \int_{t-D_i}^{\sigma} U_i(\tau)d\tau\,. \tag{164}$$

Now, note that

$$\tilde{\theta}_i(\sigma) = \tilde{\vartheta}_i(\sigma+D_i)\,, \quad \forall\sigma \in [t-D_i, t]\,. \tag{165}$$

Hence,

$$\tilde{\theta}_i(\sigma) = \tilde{\theta}_i(t - D_i) + \int_{t-D_i}^{\sigma} U_i(\tau)d\tau\,, \quad \forall \sigma \in [t - D_i, t]\,. \tag{166}$$

Applying the supremum norm to both sides of (166), we have

$$\begin{aligned}
&\sup_{t-D_i\le\sigma\le t} \left|\tilde{\theta}_i(\sigma)\right| \\
&\le \sup_{t-D_i\le\sigma\le t} \left|\tilde{\theta}_i(t-D_i)\right| + \sup_{t-D_i\le\sigma\le t} \int_{t-D_i}^{t} |U_i(\tau)|\, d\tau \\
&\le \left|\tilde{\theta}_i(t-D_i)\right| + \int_{t-D_i}^{t} |U_i(\tau)|\, d\tau \quad \text{(by Cauchy-Schwarz)} \\
&\le \left|\tilde{\theta}_i(t-D_i)\right| + \left(\int_{t-D_i}^{t} d\tau\right)^{1/2} \times \left(\int_{t-D_i}^{t} |U_i(\tau)|^2\, d\tau\right)^{1/2} \\
&\le \left|\tilde{\theta}_i(t-D_i)\right| + \sqrt{D_i}\left(\int_{t-D_i}^{t} U_i^2(\tau)d\tau\right)^{1/2}.
\end{aligned} \tag{167}$$

Now, it is easy to check

$$\left|\tilde{\theta}_i(t - D_i)\right| \le \left(\left|\tilde{\theta}_i(t - D_i)\right|^2 + \int_{t-D_i}^{t} U_i^2(\tau)d\tau\right)^{1/2}, \tag{168}$$

$$\left(\int_{t-D_i}^{t} U_i^2(\tau)d\tau\right)^{1/2} \le \left(\left|\tilde{\theta}_i(t - D_i)\right|^2 + \int_{t-D_i}^{t} U_i^2(\tau)d\tau\right)^{1/2}. \tag{169}$$

By using (168) and (169), one has

$$\begin{aligned}
&\left|\tilde{\theta}_i(t - D_i)\right| + \sqrt{D_i}\left(\int_{t-D_i}^{t} U_i^2(\tau)d\tau\right)^{1/2} \\
&\qquad \le (1 + \sqrt{D_i})\left(\left|\tilde{\theta}_i(t - D_i)\right|^2 + \int_{t-D_i}^{t} U_i^2(\tau)d\tau\right)^{1/2}.
\end{aligned} \tag{170}$$

From (167), it is straightforward to conclude that

$$\sup_{t-D_i\le\sigma\le t} \left|\tilde{\theta}_i(\sigma)\right| \le (1+\sqrt{D_i})\left(\left|\tilde{\theta}_i(t-D_i)\right|^2 + \int_{t-D_i}^{t} U_i^2(\tau)d\tau\right)^{1/2} \tag{171}$$

and, consequently,

$$\left|\tilde{\theta}_i(t)\right| \le (1+\sqrt{D_i})\left(\left|\tilde{\theta}_i(t-D_i)\right|^2 + \int_{t-D_i}^{t} U_i^2(\tau)d\tau\right)^{1/2}. \tag{172}$$

Inequality (172) can be given in terms of the periodic solution $\tilde{\theta}_i^{\Pi}(t - D_i)$, $U_i^{\Pi}(\tau)$, $\forall \tau \in$

$[t-D_i, t]$ as follows

$$\left|\tilde{\theta}_i(t)\right| \leq (1+\sqrt{D_i}) \left(\left|\tilde{\theta}_i(t-D_i) - \tilde{\theta}_i^{\Pi}(t-D_i) + \tilde{\theta}_i^{\Pi}(t-D_i)\right|^2 + \int_{t-D_i}^{t} \left[U_i(\tau) - U_i^{\Pi}(\tau) + U_i^{\Pi}(\tau)\right]^2 d\tau \right)^{1/2}. \tag{173}$$

By applying Young's inequality and some algebra, the right-hand side of (173) and $\left|\tilde{\theta}_i(t)\right|$ can be majorized by

$$\begin{aligned} \left|\tilde{\theta}_i(t)\right| \leq \sqrt{2}\ (1+\sqrt{D_i}) \Bigg(& \left|\tilde{\theta}_i(t-D_i) - \tilde{\theta}_i^{\Pi}(t-D_i)\right|^2 + \left|\tilde{\theta}_i^{\Pi}(t-D_i)\right|^2 \\ & + \int_{t-D_i}^{t} \left[U_i(\tau) - U_i^{\Pi}(\tau)\right]^2 d\tau + \int_{t-D_i}^{t} \left[U_i^{\Pi}(\tau)\right]^2 d\tau \Bigg)^{1/2}. \end{aligned} \tag{174}$$

According to the averaging theorem by [20, 22], we can conclude that the actual state converges exponentially to the periodic solution, *i.e.*, $\tilde{\theta}_i(t-D_i) - \tilde{\theta}_i^{\Pi}(t-D_i) \to 0$ and $\int_{t-D_i}^{t} \left[U_i(\tau) - U_i^{\Pi}(\tau)\right]^2 d\tau \to 0$, exponentially. Hence,

$$\limsup_{t\to+\infty} |\tilde{\theta}_i(t)| = \sqrt{2}\ \left(1+\sqrt{D_i}\right) \times \left(\left|\tilde{\theta}_i^{\Pi}(t-D_i)\right|^2 + \int_{t-D_i}^{t} [U_i^{\Pi}(\tau)]^2 d\tau \right)^{1/2}.$$

Then, from (116), we can write $\limsup_{t\to+\infty} |\tilde{\theta}(t)| = \mathcal{O}(1/\omega)$. From (12) and reminding that $\theta(t) = \hat{\theta}(t) + S(t)$ with $S(t)$ in (15), one has that $\theta(t) - \theta^* = \tilde{\theta}(t) + S(t)$. Since the first term in the right-hand side is ultimately of order $\mathcal{O}(1/\omega)$ and the second term is of order $\mathcal{O}(|a|)$, then we state (117). From (101) and (117), we get (118).

Corollary 1: *It is easy to show that the controller (114) becomes (34) in the case of output delays or equal inputs delays. Hence, the local stability/convergence results of the multiparameter Gradient ES in Figure 3 with delayed output (10) and (11), and $D \geq 0$ being simply a scalar can be directly stated for the closed-loop delayed system (26) and (34) from Theorem 2.*

3.3 Simulation Results

In this example, multivariable ES is used for finding a source of a signal (chemical, acoustic, electromagnetic, etc.) of unknown concentration field as in (101) with Hessian:

$$H = \begin{bmatrix} -2 & -2 \\ -2 & -4 \end{bmatrix}.$$

The strength of this field decays with the distance and has a local maximum at $y^* = 1$ and unknown maximizer $\theta^* = (\theta_1^*, \theta_2^*) = (0, 1)$. This is achieved without the measurement of the position vector $\theta = (\theta_1, \theta_2)$ and using only the measurement of the output scalar signal y with delay $D_{\text{out}} = 5\,\text{s}$. The two actuator paths of the vehicle are also under distinct delays $D_1^{\text{in}} = 10\,\text{s}$ and $D_2^{\text{in}} = 20\,\text{s}$. Thus, the total delays to be compensated by the predictors are $D_1 = 15\,\text{s}$ and $D_2 = 25\,\text{s}$.

The proposed schemes are slightly modified for the stated task in Figure 7 by observing that the integrator, a key adaptation element, is already present in vehicle models where the primary forces or moments acting on the vehicle are those that provide thrust/propulsion [23]. Thus, an application of our result for single and double integrators in control of autonomous vehicles modeled as point mass in the plane is shown to be possible. However, due to lack of space, we consider the simplest case of a velocity-actuated point mass only, where the additive dither in (15) is changed by $\dot{S}(t)$ since the integrator of the vehicle dynamics can be moved to the ES loop for analysis purposes. For the double integrator case, it would be needed to replace the lag filter used in (114) by lead compensator of the form $sc/(s+c)$, whose role is to recover some of the phase in feedback loop lost due to the addition of the second integrator [23].

The predictor feedback based ES controllers drive the autonomous vehicle modeled by

$$\dot{\theta}_1 = v_1\,, \qquad \dot{\theta}_2 = v_2 \tag{175}$$

to (θ_1^*, θ_2^*), whereas the ES automatically tunes v_1, v_2 to lead the vehicle to the peak of $Q(\theta)$.

Tests were performed with the following parameters: $a_1 = a_2 = 0.05$, $\omega = 5$, $\omega_1 = 7\omega$, $\omega_2 = 5\omega$, $\hat{\theta}(0) = (-1, 2)$, $K = 10^{-2}\,\text{diag}\{1,\ 0.5\}$ and $c = 20$.

Figure 8 shows the system output $y(t)$ in 3 situations: (a) free of delays, (b) in

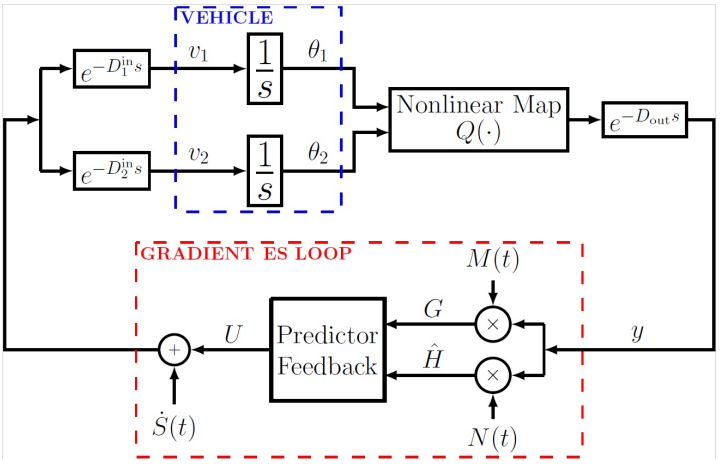

Figure 7 Source seeking under delays for velocity-actuated point mass with additive dither $\dot{S}(t) = [a_1\omega_1 \cos(\omega_1(t + D_1)) \;\; a_2\omega_2 \cos(\omega_2(t + D_2))]^T$. The signals $M(t)$ and $N(t)$ are chosen according to (16) and (37). The predictor-based controller (114) is used in the ES loop to compensate the total delay $D_1 = D_1^{\text{in}} + D_{\text{out}}$ and $D_2 = D_2^{\text{in}} + D_{\text{out}}$.

the presence of input-output delays but without any delay compensation and (c) with input-output delays and predictor based compensation.

Figure 9 presents relevant variables for ES. It is clear the remarkable evolution of the new prediction scheme in searching the maximum and the Hessian's $\hat{H}$.

For the case $n = 2$ (two control inputs), equation (114) is:

$$U_1(t) = \frac{c}{s+c}\left\{K_1^T\left(G(t) + \int_{t-D_1}^{t} \hat{H}_1 U_1(\tau)d\tau + \int_{t-D_1}^{t} \hat{H}_2 U_2(\tau - (D_2 - D_1))d\tau\right)\right\}, \tag{176}$$

$$\begin{aligned} U_2(t) = \frac{c}{s+c}\Bigg\{K_2^T\Bigg(&e^{A_1(D_2-D_1)}G(t) \\ &+ e^{A_1(D_2-D_1)}\int_{t-D_1}^{t} \hat{H}_1 U_1(\tau)d\tau \\ &+ e^{A_1(D_2-D_1)}\int_{t-D_1}^{t} \hat{H}_2 U_2(\tau - (D_2 - D_1))d\tau \\ &+ \int_{t-(D_2-D_1)}^{t} e^{A_1(t-\tau)}\hat{H}_2 U_2(\tau)d\tau\Bigg)\Bigg\}. \end{aligned} \tag{177}$$

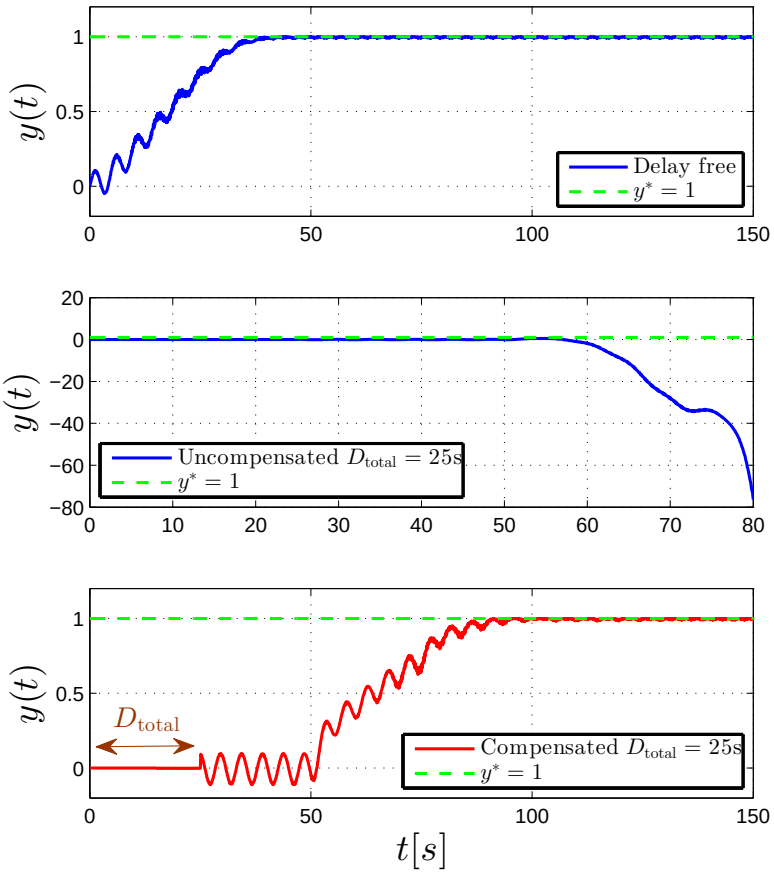

Figure 8 Time response of $y(t)$): (a) basic ES works well without delays; (b) ES goes unstable in the presence of delays ($D_{\text{total}} = D_2 = D_2^{\text{in}} + D_{\text{out}}$ is the longest delay); (c) predictor fixes this.

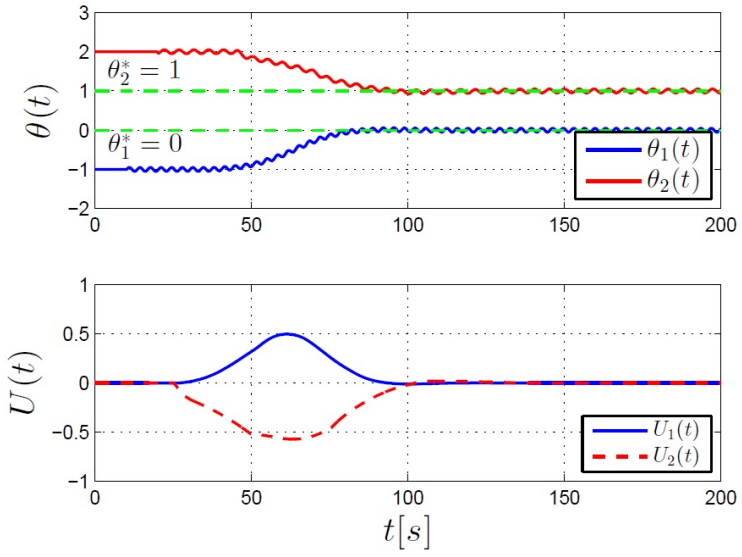

Figure 9 (a) parameter $\theta(t)$; (b) the control signal $U(t)$

4 AN ALTERNATIVE GRADIENT-BASED ES SCHEME FOR MULTIPLE AND DISTINCT INPUT DELAYS

The Extremum Seeking scheme for multiple and distinct input delays presented in Chapter 3 results in a difficult to implement controller, in which the stability analysis involve the use of backstepping transformation and averaging in infinite dimensions (3.2 - step 4). In contrast to this, the alternative gradient-based extremum seeking scheme presentend in this chapter does not require backstepping transformation, which also leads to a much more simple control feedback structure.

4.1 Alternative Gradient-Based Predictor

Throughout this chapter, $e_i \in \mathbb{R}^n$ stands for the i^{th} column of the identity matrix $I_n \in \mathbb{R}^{n\times n}$ for each $i \in \{1, 2, ..., n\}$. Without loss of generality we assume that the inputs have distinct known delays which are ordered so that

$$D = diag\{D_1, D_2, \cdots, D_n\}, \quad 0 \leq D_1 \leq \cdots \leq D_n. \tag{178}$$

Given an $\mathbb{R}^n$ -valued signal f, the notation f^D denotes

$$f^D(t) = \begin{bmatrix} f_1(t-D_1) & f_2(t-D_2) & ... & f_n(t-D_n) \end{bmatrix}^T \tag{179}$$

Let $Q : \mathbb{R}^n \to \mathbb{R}$ be a convex static map with a maximum at $\theta^* \in \mathbb{R}^n$. We assume that the point θ^* is unknown but the output of Q is available for the past input signal. More precisely, the measurable signal is given by

$$y(t) = Q(\theta^D(t)) \tag{180}$$

The purpose of the Extremum Seeking control is to estimate θ^* from the output y. To this end, define perturbation signals $S(t)$ and $M(t) \in \mathbb{R}^n$ by

$$S(t) = \begin{bmatrix} a_1 \sin(\omega_1(t+D_1)) & \cdots & a_n \sin(\omega_n(t+D_n)) \end{bmatrix}^T \tag{181}$$

$$M(t) = \begin{bmatrix} \frac{2}{a_1}\sin(\omega_1 t) & \cdots & \frac{2}{a_n}\sin(\omega_n t) \end{bmatrix}^T \tag{182}$$

The delayed signal S^D of S is a conventional perturbation signal. We also set the matrix-valued signal $N(t) \in \mathbb{R}^{n\times n}$ as

$$N_{ij}(t) = \begin{cases} \dfrac{16}{a_i^2}\left(\sin^2(\omega_i t) - \dfrac{1}{2}\right), & i = j \\ \dfrac{4}{a_i a_j}\sin(\omega_i t)\sin(\omega_j t), & i \neq j. \end{cases} \tag{183}$$

The probing frequencies ω_i's can be selected as

$$\omega_i = \omega_i' \omega = \mathcal{O}(\omega)\,, \quad i \in 1, 2, \ldots, n\,, \tag{184}$$

where ω is a positive constant and ω_i' is a rational number. One possible choice is given in [9] as

$$\omega_i' \notin \left\{\omega_j'\,,\ \frac{1}{2}(\omega_j' + \omega_k')\,,\ \omega_j' + 2\omega_k'\,,\ \omega_j' + \omega_k' \pm \omega_l'\right\}\,, \tag{185}$$

for all distinct i, j, k and l.

By using the above signals, we develop an Extremum Seeking scheme in the presence of input delays. Let the input signal be constructed as

$$\theta(t) = \hat{\theta}(t) + S(t), \tag{186}$$

where $\hat{\theta}$ is an estimate of θ^*. Then, the corresponding output signal becomes

$$y(t) = Q(\theta^D(t)) = Q\left(\hat{\theta}^D(t) + S^D(t)\right). \tag{187}$$

We introduce the estimation error $\tilde{\theta} := \hat{\theta}^D(t) - \theta^*$. Note that the error is defined with $\hat{\theta}^D$ rather than $\hat{\theta}$. With this error variable, the output signal $y(t)$ can be rewritten as

$$y(t) = Q\left(\theta^* + \tilde{\theta}(t) + S^D(t)\right)\,. \tag{188}$$

To compensate the delays, we propose the following predictor-based update law:

$$\dot{\hat{\theta}} = U(t)\,, \tag{189}$$

$$\dot{U}(t) = -cU(t) + cK\left(M(t)y(t) + N(t)y(t)\sum_{i=1}^{n} e_i \int_{t-D_i}^{t} U_i(\tau)d\tau\right) \tag{190}$$

for some positive constant $c > 0$ and diagonal matrix $K \in \mathbb{R}^{n\times n}$ with positive entries. Since $\dot{\hat{\theta}}^D(t) = U^D(t)$, differentiating the error variable $\tilde{\theta}$ with respect to t yields

$$\dot{\tilde{\theta}}(t) = U^D(t) = \sum_{i=1}^{n} e_i U_i(t - D_i), \tag{191}$$

which is in a standard form of a system with input delays. As we will see later, the terms in the parentheses in the right-hand side of (190) corresponds to a predicted value of $H\tilde{\theta}$ at some time in the future in the average sense.

4.2 Stability Analysis

Theorem 3. *Consider the control system in Figure 10 with multiple and distinct input delays according to (178) and locally quadratic nonlinear map (180). There exists $c^* > 0$ such that, $\forall c \geq c^*$, $\exists\, \omega^*(c) > 0$ such that, $\forall \omega > \omega^*$, the closed-loop delayed system (190) and (191) with state $\tilde{\theta}_i(t - D_i)$, $U_i(\tau)$, $\forall \tau \in [t - D_i, t]$ and $\forall i \in 1, 2, \ldots, n$, has a unique locally exponentially stable periodic solution in t of period Π, denoted by $\tilde{\theta}_i^{\Pi}(t-D_i)$, $U_i^{\Pi}(\tau)$, $\forall \tau \in [t - D_i, t]$ and $\forall i \in \{1, 2, \ldots, n\}$, satisfying, $\forall t \geq 0$:*

$$\left(\sum_{i=1}^{n} \left[\tilde{\theta}_i^{\Pi}(t - D_i)\right]^2 + \left[U_i^{\Pi}(t)\right]^2 + \int_{t-D_i}^{t} \left[U_i^{\Pi}(\tau)\right]^2 d\tau\right)^{1/2} \leq \mathcal{O}(1/\omega)\,. \tag{192}$$

Furthermore,

$$\limsup_{t\to+\infty} |\theta(t) - \theta^*| = \mathcal{O}(|a| + 1/\omega)\,, \tag{193}$$

$$\limsup_{t\to+\infty} |y(t) - y^*| = \mathcal{O}(|a|^2 + 1/\omega^2)\,, \tag{194}$$

where $a = [a_1\ a_2\ \cdots\ a_n]^T$.

Proof. A PDE representation of the closed-loop system (190), (191) is given by

$$\dot{\tilde{\theta}}(t) = u(0, t), \tag{195}$$

$$u_t(x,t) = D^{-1}u_x(x,t), \quad x \in (0,1), \tag{196}$$

$$u(1,t) = U(t) \tag{197}$$

$$\dot{U}(t) = -cU(t) + cK\left(M(t)y(t) + N(t)y(t)\int_0^1 Du(x,t)dx\right), \tag{198}$$

where $u(x,t) = (u_1(x,t), u_2(x,t), \cdots, u_n(x,t))^T \in \mathbb{R}^n$. It is easy to see that the solution of (196) under the condition (197) is represented as $u_i(x,t) = U_i(D_i x + t - D_i)$ for each $i \in \{1, 2, \cdots, n\}$. Hence, we have

$$\int_0^1 D_i u_i(x,t)dx = \int_{t-D_i}^t u_i(\frac{\tau - t + D_i}{D_i}, t)d\tau = \int_{t-D_i}^t U_i(\tau)d\tau. \tag{199}$$

This means that

$$\int_0^1 Du(x,t)dx = \sum_{i=1}^n e_i \int_0^1 D_i u_i(x,t)dx = \sum_{i=1}^n e_i \int_{t-D_i}^t U_i(\tau)d\tau. \tag{200}$$

Thus, we can recover (191) from (196)-(198). The average system associated with (195)-(198) is given by

$$\dot{\tilde{\theta}}_{av}(t) = u_{av}(0,t), \tag{201}$$

$$u_{av,t}(x,t) = D^{-1}u_{av,x}(x,t), \quad x \in (0,1) \tag{202}$$

$$u_{av}(1,t) = U_{av}(t) \tag{203}$$

$$\dot{U}_{av}(t) = -cU_{av}(t) + cKH\left(\tilde{\theta}_{av}(t) + \int_0^1 Du_{av}(x,t)dx\right), \tag{204}$$

where we have used the fact that the averages of $M(t)y(t)$ and $N(t)y(t)$ are calculated as $H\tilde{\theta}_{av}(t)$ and H. For simplicity of notation, let us introduce the following auxiliary variables

$$\vartheta(t) := H\left(\tilde{\theta}_{av}(t) + \int_0^1 Du_{av}(x,t)dx\right), \tag{205}$$

$$\tilde{U} = U_{av} - K\vartheta. \tag{206}$$

With this notation, (204) can be represented simply as $\dot{U}_{av} = -c\tilde{U}$. In addition, differentiating (205) with respect to t yields

$$\dot{\vartheta} = HU_{av}(t). \tag{207}$$

We prove the exponential stability of the closed-loop system by using the Lyapunov functional defined by

$$V(t) = \vartheta(t)^T K\vartheta(t) + \frac{1}{4}\lambda_{\min}(-H)\int_0^1 (1+x)u_{av}(x,t)^T Du_{av}(x,t)dx + \frac{1}{2}\tilde{U}(t)^T(-H)\tilde{U}(t). \tag{208}$$

Recall that K and D are diagonal matrices with positive entries and that H is a negative-definite matrix. Hence, all of K, D, and $-H$ are positive-definite matrices. For simplicity of notation, we omit explicit dependence of variables on t. The time derivative of V is given by

$$\begin{aligned}
\dot{V} &= 2\vartheta^T KHU_{av} + \frac{1}{2}\lambda_{\min}(-H)U_{av}^T U_{av} - \frac{1}{4}\lambda_{\min}(-H)u(0)^T u(0) \\
&\quad - \frac{1}{4}\lambda_{\min}(-H)\int_0^1 u_{av}(x)^T u_{av}(x)dx + \tilde{U}^T(-H)\left(\dot{U}_{av} - KHU_{av}\right) \\
&\leq 2\vartheta^T KHU_{av} + \frac{1}{2}U_{av}^T(-H)U_{av} \\
&\quad - \frac{1}{8D_{\max}}\lambda_{\min}(-H))\int_0^1 (1+x)u_{av}(x)^T Du_{av}(x)dx \\
&\quad + \tilde{U}^T(-H)\dot{U}_{av} + \tilde{U}^T(-H)K(-H)U_{av}.
\end{aligned} \tag{209}$$

Applying Young's inequality to the last term leads to

$$\tilde{U}^T(-H)K(-H)U_{av} \leq \frac{1}{2}\tilde{U}^T(-HKHKH)\tilde{U} + \frac{1}{2}U_{av}^T(-H)U_{av}. \tag{210}$$

Then, completing the square yields

$$\begin{aligned}\dot{V} \leq\ & \tilde{U}^T(-H)\tilde{U} - \vartheta^T K(-H)K\vartheta \\ & - \frac{1}{8D_{\max}}\lambda_{\min}(-H)\int_0^1 (1+x)u_{av}(x)^T Du_{av}(x)dx \\ & + \tilde{U}^T(-H)\dot{U}_{av} + \frac{1}{2}\tilde{U}^T(-HKHKH)\tilde{U} \\ =\ & \tilde{U}^T(-H)\left(\dot{U}_{av} + c^*\tilde{U}\right) - \vartheta^T K(-H)K\vartheta \\ & - \frac{1}{8D_{\max}}\lambda_{\min}(-H)\int_0^1 (1+x)u_{av}(x)^T Du_{av}du, \end{aligned} \tag{211}$$

where $c^* := 1 + \lambda_{\max}(-HKHKH)/\lambda_{\min}(-H)$. Hence, by setting $\dot{U}_{av} = -c\tilde{U}$ for some $c > c^*$, we see that there exists $\mu > 0$ such that $\dot{V} \leq \mu V$. Finally, it is not difficult to find positive constants $\alpha, \beta > 0$ such that

$$\alpha\left(|\tilde{\theta}_{av}(t)|^2 + \int_0^1 |u(x,t)|^2 dx + |\tilde{U}_{av}(t)|^2\right) \leq V(t) \leq \beta\left(|\tilde{\theta}_{av}(t)|^2 + \int_0^1 |u(x,t)|^2 dx + |\tilde{U}_{av}(t)|^2\right). \tag{212}$$

Therefore, the average system (201)-(204) is exponentially stable as long as $c > c^*$. The remaining steps of the proof follow exactly the steps 6 and 7 of chapter 3.

4.3 Simulation Results

In order to evaluate the multidimentional version of the delay-compensated extremum seeking, we consider the following static quadratic map:

$$Q(\theta) = 1 + \frac{1}{2}\left(2(\theta_1)^2 + 4(\theta_2 - 1)^2 + 4\theta_1(\theta_2 - 1)\right), \tag{213}$$

subject to an input delay of $D = (35, 40)$. The extremum points are $\theta^* = (0, 1)$ and $y^* = 1$, and the Hessian of the map is

$$H = -\begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}. \tag{214}$$

For this case ($n = 2$), the predictor controller equation is given by

$$\dot{U}(t) = -cU(t) + c\begin{bmatrix} K_1 & 0 \\ 0 & K_2 \end{bmatrix}\Bigg(M(t)y(t) + N(t)y(t)\Bigg(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\int_{t-D_1}^{t} U_1(\tau)d\tau + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\int_{t-D_2}^{t} U_2(\tau)d\tau\Bigg)\Bigg). \tag{215}$$

Since $M(t)$ and $N(t)$ are

$$M(t) = \begin{bmatrix} M_1(t) \\ M_2(t) \end{bmatrix} = \begin{bmatrix} \frac{2}{a_1}sin(\omega_1 t) \\ \frac{2}{a_2}sin(\omega_2 t) \end{bmatrix} \tag{216}$$

$$N(t) = \begin{bmatrix} N_{11}(t) & N_{12}(t) \\ N_{21}(t) & N_{22}(t) \end{bmatrix} = \begin{bmatrix} \frac{16}{a_1^2}\left(sin^2(\omega_1 t) - \frac{1}{2}\right) & \frac{4}{a_1 a_2}sin(\omega_1 t)sin(\omega_2 t) \\ \frac{4}{a_2 a_1}sin(\omega_2 t)sin(\omega_1 t) & \frac{16}{a_2^2}\left(sin^2(\omega_2 t) - \frac{1}{2}\right) \end{bmatrix}, \tag{217}$$

The predictor's equation (215) can be written as $U_1(t)$ and $U_2(t)$ as follows:

$$U_1(t) = \frac{cK_1}{s+c}\Bigg\{\hat{G}(t) + \hat{H}_1\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\int_{t-D_1}^{t} U_1(\tau)d\tau + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\int_{t-D_2}^{t} U_2(\tau)d\tau\right)\Bigg\} \tag{218}$$

$$U_2(t) = \frac{cK_2}{s+c}\Bigg\{\hat{G}(t) + \hat{H}_2\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\int_{t-D_1}^{t} U_1(\tau)d\tau + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\int_{t-D_2}^{t} U_2(\tau)d\tau\right)\Bigg\} \tag{219}$$

When comparing (218) and (219) with the predictor's equations (176) and (177) in the $n = 2$ example in Chapter 3, it is possible to notice that the new algorithm results in a simplified controller with easier implementation.

The proposed alternative gradient-based ES under multiple-input delays is shown in the block diagram of Figure 10. For this $n = 2$ example, $D = diag\{D_1, D_2\}$ and $K = diag\{K_1, K_2\}$. The predictor feedback is implemented according to (218) and (219), the additive dither signal is $S(t) = \begin{bmatrix} a_1 sin(\omega_1(t + D_1)) & a_2 sin(\omega_2(t + D_2)) \end{bmatrix}^T$ and $M(t)$ and $N(t)$ are given by (216) and (217) respectively.

In our simulations, we use low-pass and washout filters with corner frequencies ω_h

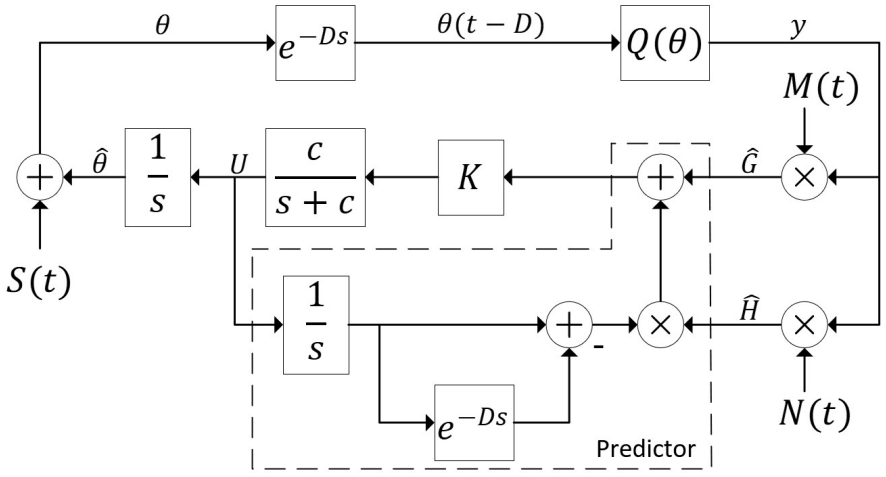

Figure 10 Block diagram of the prediction scheme for multiple-input delay compensation. $\hat{G}$ and $\hat{H}$ are, respectively, the Gradient and Hessian estimates. $D = diag\{D_1, D_2\}$ and $K = diag\{K_1, K_2\}$

and ω_l to improve the controller performance as usual in extremum seeking designs.

In what follows, we present numerical simulations of the predictor (215). We performed our tests with $c = 20$, $K_1 = \frac{1}{100}$, $K_2 = \frac{1}{200}$, $a_1 = a_2 = 0.05$, $\omega = 0.5$, $\omega_1 = 17.5\omega$, $\omega_2 = 12.5\omega$, $\omega_h = \omega_l = \frac{\omega}{5}$, and $\hat{\theta}(0) = (1, 1)$.

Figure 11 shows the system output $y(t)$ in 3 situations: (a) free of input delays, (b) in the presence of multiple and distinct input-delays without any delay compensation and (c) with input-delays and predictor-based compensation.

Figure 12 presents relevant variables for ES: (a) the time response of the parameter $\theta(t)$, (b) the control signal $U(t)$ and (c) the Hessian's estimate $\hat{H}$. It is clear the remarkable evolution of the new prediction scheme in searching the maximum and the Hessian's parameters of H^{-1}. This ultimate exact perturbation based estimation allows the perfect delay compensation.

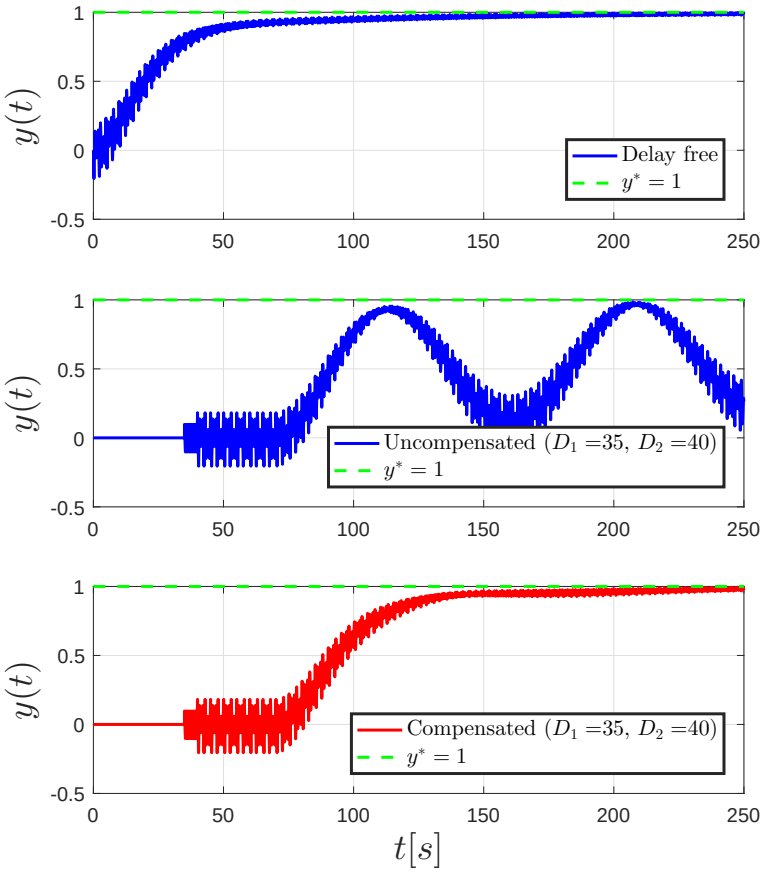

Figure 11 Alternative gradient-based ES under input-output delays (time response of $y(t)$: (a) basic ES works well without delays; (b) ES goes unstable in the presence of delays; (c) predictor fixes this.

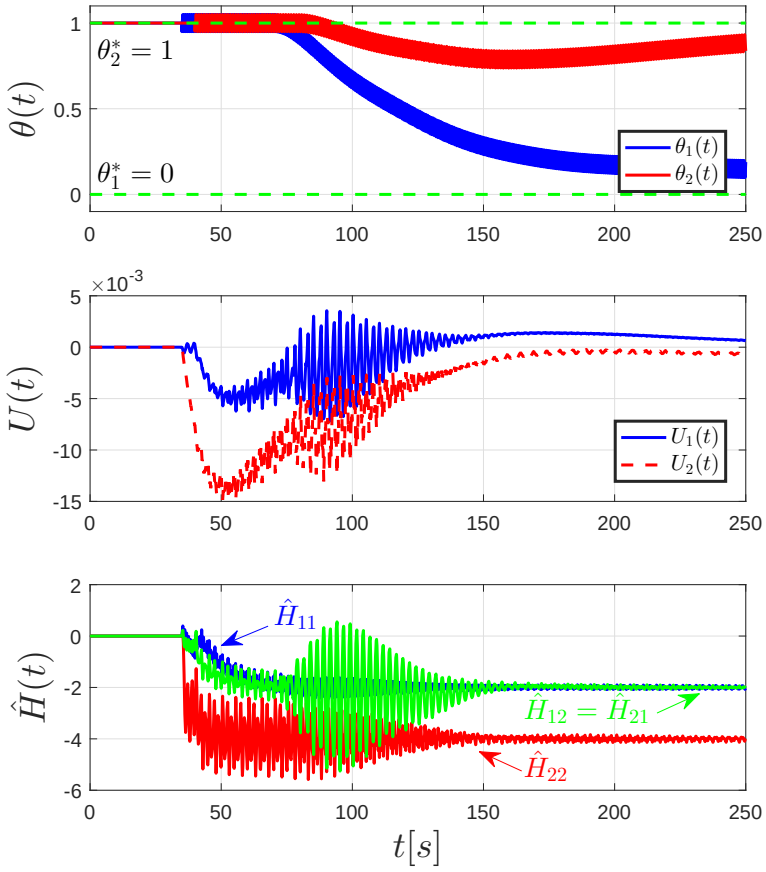

Figure 12 Alternative gradient-based ES under multiple input-delays: (a) parameter $\theta(t)$; (b) the control signal $U(t)$; (c) Hessian's estimate $\hat{H}(t)$. The elements of $\hat{H}(t)$ converge to the unknown elements of $H(t)$.

CONCLUSION

A new gradient extremum seeking controller was developed for multiparameter real-time optimization in the presence of distinct actuator delays. The control scheme introduced here for delay compensation uses prediction feedback with perturbation-based estimate of the Hessian associated with an adequate tune of the dither signals. The novel predictor feedback is much simpler than the previous gradient case found in the literature. The generalization for multi-input-single-output maps with distinct input delays has to perform prediction of the cross-coupling of the channels. As a further advance in this challenging scenario is that the contributions are achieved without invoking the backstepping methodology usually employed in the publications concerning extremum seeking under delays.

As future work, suggestion to continue developing extremum seeking theory with a stochastic version of the proposed algorithm, application to dynamic maps rather than static ones, and unknown and distributed delays.

REFERENCES

[1] TSUBAKINO, D. *An alternative gradient-based extremum seeking scheme.* [S.l.], August 2017.

[2] KRSTIĆ, M.; WANG, H. H. Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, v. 36, p. 595–601, 2000.

[3] ARIYUR, K. B.; KRSTIĆ, M. *Real-time Optimization by Extremum Seeking Control.* [S.l.]: John Wiley & Sons, 2003.

[4] KRSTIĆ, M. Extremum seeking control. In: T. Samad and J. Baillieul (Ed.). *Encyclopedia of Systems and Control.* Berlin, Germany: Springer, 2014.

[5] GUAY, M.; ZHANG, T. Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties. *Automatica*, v. 39, p. 1283–1293, 2003.

[6] ADETOLA, V.; GUAY, M. Parameter convergence in adaptive extremum-seeking control. *Automatica*, v. 43, n. 1, p. 105–110, 2007.

[7] TAN, Y.; NESIĆ, D.; MAREELS, I. On non-local stability properties of extremum seeking control. *Automatica*, v. 42, p. 889–903, 2006.

[8] TANA, Y. et al. On global extremum seeking in the presence of local extrema. *Automatica*, v. 45, p. 245–251, 2009.

[9] GHAFFARI, A.; KRSTIĆ, M.; NESIĆ, D. Multivariable Newton-based extremum seeking. *Automatica*, v. 48, p. 1759–1767, 2012.

[10] GUAY, M.; DOCHAIN, D. A time-varying extremum-seeking control approach. *Automatica*, v. 51, p. 356–363, 2015.

[11] NESIĆ, D.; MOHAMMADI, A.; MANZIE, C. A framework for extremum seeking control of systems with parameter uncertainties. *IEEE Transactions on Automatic Control*, v. 58, n. 2, p. 435–448, 2013.

[12] LIU, S.-J.; KRSTIĆ, M. *Stochastic Averaging and Stochastic Extremum Seeking.* [S.l.]: Springer, 2012.

[13] KRSTIĆ, M. Delay compensation for nonlinear, adaptive, and PDE systems. *American Control Conference*, Birkhaüser, 2009.

[14] OLIVEIRA, T. R.; KRSTIĆ, M.; TSUBAKINO, D. Multiparameter extremum seeking with output delays. *American Control Conference*, Chicago, p. 152–158, 2015.

[15] NESIĆ, D. et al. A unifying approach to extremum seeking: adaptive schemes based on the estimation of derivatives. In: *IEEE Conference on Decision and Control*. Atlanta: [s.n.], 2010. p. 4625–4630.

[16] OLIVEIRA, T. R.; KRSTIĆ, M.; TSUBAKINO, D. Extremum seeking for static maps with delays. *IEEE Transactions on Automatic Control*, v. 62, n. 5, p. 1911–1926, April 2017.

[17] TSUBAKINO, D.; OLIVEIRA, T. R.; KRSTIĆ, M. Predictor-feedback for multi-input LTI systems with distinct delays. *American Control Conference*, Chicago, p. 571–576, 2015.

[18] MALISOFF, M.; KRSTIĆ, M. Multivariable extremum seeking with distinct delays using a one-stage sequential predictor. *American Control Conference*, accepted paper, Denver, 2020.

[19] KHALIL, H. K. *Nonlinear Systems*. [S.l.]: Prentice-Hall, 1996.

[20] HALE, J. K.; LUNEL, S. M. V. Averaging in infinite dimensions. *Journal of Integral Equations and Applications*, v. 2, p. 463–494, 1990.

[21] KRSTIĆ, M. Lyapunov tools for predictor feedbacks for delay systems inverse optimality and robustness to delay mismatch. *Automatica*, v. 44, p. 2930–2935, 2008.

[22] LEHMAN, B. The influence of delays when averaging slow and fast oscillating systems: overview. *IMA Journal of Mathematical Control and Information*, v. 19, p. 201–215, 2002.

[23] ZHANG, C.; SIRANOSIAN, A.; KRSTIĆ, M. Extremum seeking for moderately unstable systems and for autonomous vehicle target tracking without position measurements. *Automatica*, v. 43, p. 1832–1839, 2007.