Fast Extremum Seeking Control for a Class of Generalized Hammerstein Systems with the Knowledge of Relative Degree

Hengchang Liu[1], Ying Tan[1], Tomislav Bačeк[1], Dana Kulić[2], Denny Oetomo[1] and Chris Manzie[1]

***Abstract*— This work extends the existing fast extremum seeking control (ESC) for a class of Hammerstein systems to a class of generalized Hammerstein systems, in which the nonlinear affine dynamic system is connected directly after a given cost function. With the relative degree information of the unknown nonlinear dynamics, a new output is generated. The mapping between the new output and the input has two parts. The first part is proportional to the cost function and the second part is related to the state. By inserting a fast dither signal, the proposed ESC can seek the optimum of this cost function without time-scale separation. Our main results show that with proper selection of tuning parameters, this scheme can achieve arbitrarily fast semi-global practical asymptotic (SPA) convergence. Simulation results support the theoretical findings.**

I. Introduction

Extremum seeking control (ESC) has been developed to seek the optimum of the mapping between the input and steady-state output of a dynamical system without the knowledge of the system. Over almost one hundred years of developments [5, 2, 13, 19, 12], ES algorithms have been used in various applications such as robotics [20], magnetic levitation systems [4], renewable energy [9], flight formation [3], and so on.

In principle, ESC seeks an optimal value for an unknown cost function. In particular, perturbation-based ESC utilizes dither signals to perturb the input to estimate the gradient, leading to the convergence. When a dynamic system is considered, the singular perturbation technique [11, Chapter 11] has been widely used to obtain the input-to-steady-state-output mapping. By tuning the ESC loop sufficiently slower than the dynamics, in a slower time-scale, the dynamics will die out quickly so that the input-to-steady-state-output mapping will be dominating. Hence, the convergence of the ESC has to be tuned slow compared with the dynamics of the system, leading to possibly slow convergence when the system dynamics are slow.

[1] School of Electrical, Mechanical and Infrastructure Engineering, The University of Melbourne, Parkville, VIC 3010, Australia hengchangl@student.unimelb.edu.au, {yingt,tomislav.bacek,doetomo,manziec}@unimelb.edu.au

[2] Department of Mechanical and Aerospace Engineering, Monash University, Clayton, VIC 3168, Australia dana.kulic@monash.edu

Many algorithms have been proposed to attempt to remove the above mentioned time-scale separation so that fast convergence can be achieved. Moase et.al proposed the fast ESC for a class of Hammerstein systems [15], which have a nonlinear cost function followed by a linear time-invariant (LTI) dynamic system. This idea has been extended to a class of Wiener-Hammerstein systems [14] where two LTI models are connected via a nonlinear mapping, and [17] provided the discrete version of [14]. Furthermore, proportional-integral ESC is another type of ESC scheme that remove the time-scale separation for a class of Wiener-type systems [7, 8].

As the dynamics considered in [15] are only LTI, this limits the applicability of the fast ESC. This work extends the results obtained in [15] to a more general setting when the nonlinear cost function is followed by a nonlinear dynamics. Such a system is called the generalized Hammerstein system.

Similar to [15], the relative degree information of the nonlinear dynamics in the generalized Hammerstein model is used to generate a new output. The relationship between this new output signal and the input signal is dominated by a static mapping, perturbed by unknown system dynamics. The static mapping is related to the cost function and has the same optimal point. By injecting high frequency dither signals, the system dynamics can be treated as slowly time-varying variables. Consequently, the effect of the unknown perturbation from the system can be ignored by using averaging technique, making the proposed fast ESC working. Our first result (Theorem 1) shows how the fast ESC can seek the optimum of a cost function without time-scale separation. Moreover, this proposed fast ESC has an extra design freedom to achieve any preferred convergence speed. Corollary 1 extends Theorem 1 to the case when the cost function has multiple inputs.

This paper is organized as follows. Section 2 presents preliminaries and problem formulation. Section 3 shows the main results. An extension to multi-variable case is covered in Section 4. Simulation examples validate the obtained results in Section 5.

II. Preliminaries and Problem Formulation

A. Preliminaries

The notation $\mathcal{R}$ represents the set of all real numbers. For any vector $\mathbf{x} \in \mathcal{R}^n$, $|\mathbf{x}|$ represents its Euclidean norm,

which is defined as $|\mathbf{x}| \triangleq \sqrt{\mathbf{x}^\intercal \mathbf{x}}$, where $(\cdot)^\intercal$ represents the transpose. A continuous function $\alpha : \mathcal{R}_{\geq 0} \to \mathcal{R}_{\geq 0}$ is said to be of class $\mathcal{K}$ if it is zero at zero and strictly increasing. And it is said to belong yo class $\mathcal{K}_\infty$ if $\alpha(r) \to \infty$ as $r \to \infty$. A continuous function $\sigma : \mathcal{R}_{\geq 0} \to \mathcal{R}_{\geq 0}$ is said to be of class $\mathcal{L}$ if it is converging to zero as its argument grows unbounded. A continuous function $\beta : \mathcal{R}_{\geq 0} \times \mathcal{R}_{\geq 0} \to \mathcal{R}_{\geq 0}$ is said to belong to class $\mathcal{KL}$ if, for each fixed s, the mapping $\beta(r, s)$ belongs to class $\mathcal{K}$ with respect to r and for each fixed r, the mapping $\beta(r, s)$ is decreasing with respect to s and $\beta(r, s) \to 0$ as $s \to \infty$ [11]. The gradient and Hessian matrix of a sufficiently smooth function $F : \mathcal{R}^m \to \mathcal{R}$ are denoted by ∇F and $\nabla^2 F$ respectively. A function $\delta_1(\varepsilon) = O(\delta_2(\varepsilon))$ if there exist positive constants k and c such that $|\delta_1(\varepsilon)| \leq k\,|\delta_2(\varepsilon)|$, $\forall\, |\varepsilon| < c$ [11].

B. Problem formulation

We first start from a class of unknown single-input-single-output (SISO) nonlinear dynamical systems:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u, \\ y &= h(\mathbf{x}), \end{aligned} \tag{1}$$

where $u \in \mathcal{R}$, $\mathbf{x} \in \mathcal{R}^n$, $y \in \mathcal{R}$. The unknown nonlinear mappings $\mathbf{f} : \mathcal{R}^n \to \mathcal{R}^n$, $\mathbf{g} : \mathcal{R}^n \to \mathcal{R}^n$, $h : \mathcal{R}^n \to \mathcal{R}$ are continuously differentiable. The Lie derivative of $h(\cdot)$ along $\mathbf{f}(\cdot)$ is denoted as $L_\mathbf{f} h(\mathbf{x}) = \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x})$. The Lie derivative of $h(\cdot)$ along $\mathbf{g}(\cdot)$ is denoted as $L_\mathbf{g} h(\mathbf{x}) = \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} \mathbf{g}(\mathbf{x})$.

The nonlinear system (1) satisfies the following assumptions. The first assumption assumes the knowledge of the relative degree of the system (1) though nonlinear mappings $\mathbf{f}(\cdot)$, $\mathbf{g}(\cdot)$, and $h(\cdot)$ are unknown. This assumption plays a key role in achieving fast ESC.

Assumption 1: The nonlinear system (1) has a global relative degree[1] 1, which indicates that $L_\mathbf{g} h(\mathbf{x}) \neq 0$ for any $\mathbf{x} \in \mathcal{R}^n$. □

Without losing generosity, we further assume $L_\mathbf{g} h(\mathbf{x}) > 0$ for any $\mathbf{x} \in R^n$. It is noted that the similar design principle can be used for nonlinear systems (1) with an arbitrary relative degree.

The second assumption assumes the parameterized stabilization for system (1), which can be found in [19, Assumptions 1-2].

Assumption 2: There exists a smooth function $\mathbf{l} : \mathcal{R} \to \mathcal{R}^n$ such that

$$\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u = 0 \ \textit{if and only if}\ \mathbf{x} = \mathbf{l}(u). \tag{2}$$

Moreover, for each $u \in \mathcal{R}$, the equilibrium $\mathbf{x} = \mathbf{l}(u)$ of system (1) is globally asymptotically stable, uniformly in u. □

[1]The definition of relative degree and global relative degree can be found in [10, Chapter 4].

Next we assume that there exists a unknown cost function $F(\cdot) : \mathcal{R} \to \mathcal{R}$, which satisfies the following assumption.

Assumption 3: There exists a unique $\theta^* \in \mathcal{R}$ that maximises $F(\cdot)$, and the following conditions hold:

$$\begin{aligned} &\frac{dF(\theta)}{d\theta}\Big|_{\theta=\theta^*} = 0, \frac{d^2F(\theta)}{d\theta^2}\Big|_{\theta=\theta^*} < 0 \\ &\frac{dF(\sigma+\theta^*)}{d\theta}\sigma < -\alpha_q(|\sigma|), \qquad \forall \sigma \in \mathcal{R}, \end{aligned} \tag{3}$$

where $\alpha_q(\cdot) \in \mathcal{K}_\infty$. □

Assumption 3 has been widely used in ESC (see [19] for example). By connecting the cost function $F(\cdot)$ and the nonlinear system (1) in series, it leads to the generalized Hammerstein model

$$\begin{aligned} u &= F(\theta), \\ \dot{\mathbf{x}} &= \mathbf{f}(x) + \mathbf{g}(x)u, \\ y &= h(\mathbf{x}). \end{aligned} \tag{4}$$

If the dynamics in the system (4) are LTI, i.e., $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$, $\mathbf{g}(\mathbf{x}) = B$, and $h(\mathbf{x}) = C\mathbf{x}$ for matrices (A, B, C) with appropriate dimensions, the dynamic system (4) is a Hammerstein system as discussed in [15].

III. Fast ESC for a SISO System

The diagram of the fast ESC for the system (4) is shown in Fig.1. This scheme consists two parts: a non-causal differentiator and a simplest perturbation-based ESC used in [19] with a simple modification by using $b = \frac{1}{a}$ so that the convergence speed of ESC does not depend on the choice of a.

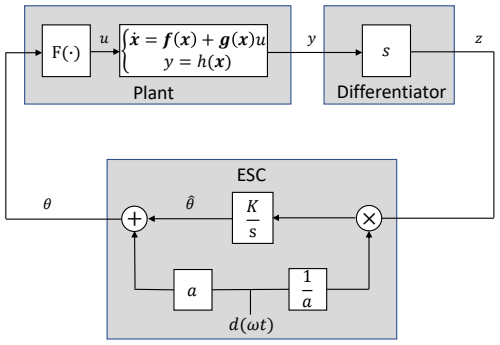

Fig. 1. Block diagram of the fast ESC for the SISO generalized Hammerstein model

Remark 1: The proposed fast ESC shares the similar design idea of the fast ESC proposed for the Hammerstein model presented in [15] with simpler structure and proof. It is highlighted that compared with LTI systems in [15], nonlinear dynamics considered in this work are more general.

Remark 2: It is noted that a differentiator is used in this diagram for the simplicity of presentation. It is non-causal and will magnify high frequency noises. In engineering applications, filtered differentiators have been widely used,

i.e., $\frac{s}{\tau s+1}$ can replace s for some $\tau > 0$ (similar techniques are used in [15]). How to choose the cut-off frequency $\frac{1}{\tau}$ is application-driven. In our simulations, we will show how this cut-off frequency will affect the performance of the proposed fast ESC.

In this diagram, the dither signal takes the form as $d(\omega t) = \sin(\omega t)$. The close-loop system in Fig.1 becomes

$$\begin{aligned}
&\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})F(\hat{\theta} + a\sin(\omega t)), \mathbf{x}(t_0) \in \mathcal{R}^n,\\
&y = h(\mathbf{x}),\\
&z = \dot{y},\\
&\dot{\hat{\theta}} = \frac{K}{a} z\sin(\omega t), \qquad \hat{\theta}(t_0) \in \mathcal{R},
\end{aligned} \tag{5}$$

where $(a, \omega, K) \in \mathcal{R}^3_{>0}$ are the tuning parameters. a is small, while K and ω are large parameters. Parameters a, K, ω play important roles in achieving certain necessary time-scale separation, so that fast convergence of the ESC can be achieved.

Let $\mathbf{x}^* = \mathbf{l}(F(\theta^*))$, where $\mathbf{l}$ is from Assumption 2 and θ^* is from Assumption 3, then we can introduce a new coordinate: $\tilde{\mathbf{x}} = \mathbf{x} - \mathbf{x}^*$, $\tilde{\theta} = \hat{\theta} - \theta^*$. The system (5) in the new coordinate takes the following form:

$$\begin{aligned}
\dot{\tilde{\mathbf{x}}} =& \mathbf{f}(\tilde{\mathbf{x}} + \mathbf{x}^*) + \mathbf{g}(\tilde{\mathbf{x}} + \mathbf{x}^*)F\left(\tilde{\theta} + \theta^* + a\sin(\omega t)\right),\\
\dot{\tilde{\theta}} =& \frac{K}{a}\left(L_{\mathbf{f}}h + L_{\mathbf{g}}hF(\tilde{\theta} + \theta^* + a\sin(\omega t)\right)\sin(\omega t),
\end{aligned} \tag{6}$$

for all $\tilde{\mathbf{x}}(t_0) \in \mathcal{R}^n$, $\tilde{\theta}(t_0) \in \mathcal{R}$, and $t \geq t_0 \geq 0$.

Let $\varepsilon = \frac{K}{a\omega}$. The convergence properties of the fast ESC proposed in Fig.1 are summarized Theorem 1.

Theorem 1: Suppose the Assumptions $1-3$ hold for the system (6). Given each strictly positive triplet (Δ, ν, K), there exist $\beta_1, \beta_2 \in \mathcal{KL}$ and a positive constant $a^* > 0$ such that for any $a \in (0, a^*)$, there exists $\varepsilon^* > 0$ such that for any $\varepsilon \in (0, \varepsilon^*)$, the solutions of the system (6) satisfy:

$$\begin{aligned}
&|\tilde{\mathbf{x}}(t)| \leq \beta_1\left(|\tilde{\mathbf{x}}(t_0)|, (t - t_0)\right) + \nu,\\
&\left|\tilde{\theta}(t)\right| \leq \beta_2\left(\left|\tilde{\theta}(t_0)\right|, K(t - t_0)\right) + \nu,
\end{aligned} \tag{7}$$

for all $t \geq t_0 \geq 0$ and $\left\|\begin{bmatrix} \tilde{\mathbf{x}}(t_0) \\ \tilde{\theta}(t_0) \end{bmatrix}\right\| \leq \Delta$.

The proof of Theorem 1 is presented in Appendix.

Remark 3: Theorem 1 shows the system (6) can achieve semi-global practical asymptotic (SPA) stability. This means that the trajectories $\tilde{\mathbf{x}}(t)$ and $\tilde{\theta}(t)$ can converge to an arbitrarily small neighborhood of the origin (a ball of radius ν centered at the origin) from any given initial condition starting from a compact set Δ by tuning parameters a sufficiently small and ε sufficiently small (dither frequency ω sufficiently large). Such SPA property was shown in [19, Theorem 1]. However the classic ESC in [19] uses slow dither signals to guarantee the necessary time-scale separation, and the convergence speed has to be slow as it is proportional to $a^2\delta\omega$ for small a, δ, and ω. In this work, the convergence speed of $\tilde{\theta}$ is only dependent on the tuning parameter K (see (7)), which is a predefined parameter. By increasing the parameter K, the convergence speed of the proposed fast ESC can increase.

Remark 4: It is highlighted that in the proof, only the averaging techniques are applied without using singular perturbation technique. This indicates that there is no need to separate the time-scale between the nonlinear dynamics (1) and the updating law ($\hat{\theta}$-dynamics in (6)), resulting to a fast convergence speed. Details can be found in the proof of Theorem 1.

IV. Extension to Cost Function with Multiple Inputs

This section extends the fast ESC proposed in Fig.1 to a more general case: a multi-input cost function $F(\cdot) : \mathcal{R}^m \to \mathcal{R}$, which satisfies the following assumption:

Assumption 4: There exists a unique $\boldsymbol{\theta}^* \in \mathcal{R}^m$ that maximises $F(\cdot)$, and the following holds:

$$\begin{aligned}
&\nabla F(\boldsymbol{\theta}^*) = \mathbf{0}, \nabla^2 F(\boldsymbol{\theta}^*) < 0,\\
&\nabla F^T(\boldsymbol{\sigma} + \boldsymbol{\theta}^*)\boldsymbol{\sigma} < -\alpha_Q(|\boldsymbol{\sigma}|), \forall \boldsymbol{\sigma} \in \mathcal{R}^m,
\end{aligned} \tag{8}$$

where $\alpha_Q(\cdot) \in \mathcal{K}_\infty$. □

Assumption 4 is an extension of Assumption 3 to a multiple input case. This leads to the following diagram shown in Fig.2.

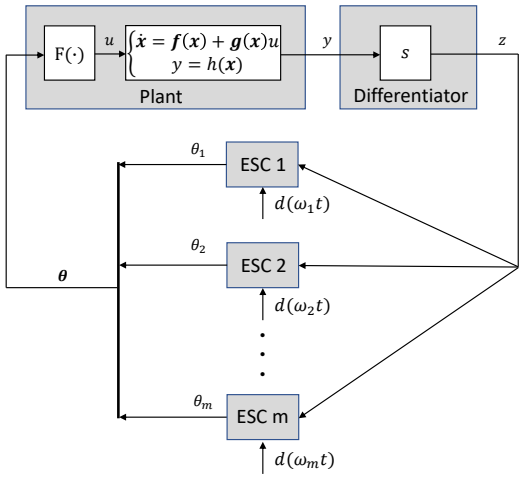

Fig. 2. Block diagram of the Fast ES algorithm for the MISO nonlinear Hammerstein model

The dynamics in Fig.2:

$$\begin{aligned}
u &= F(\boldsymbol{\theta}),\\
\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u,\\
y &= h(\mathbf{x}),
\end{aligned} \tag{9}$$

where $\mathbf{x} \in \mathcal{R}^n$, $u \in \mathcal{R}$, $\boldsymbol{\theta} = [\theta_1 \dots \theta_m]^T \in \mathcal{R}^m$ and $y \in \mathcal{R}$. This structure is similar to a multivariable ESC setting as discussed in [6], requiring a vector dither signal to be rich

enough to learn unknown optimal vector $\boldsymbol{\theta}^*$. The following condition is used to generate the needed PE condition as discussed in [6].

Condition 1: The dither signal $\mathbf{d}(\omega t)$ is selected as

$$\begin{aligned}\mathbf{d}(\omega t) =&[d(\omega t_1) \ \cdots \ d(\omega t_m)]^T\\ =&[\sin(\omega_1 t) \ \cdots \ \sin(\omega_m t)]^T,\end{aligned} \tag{10}$$

where $\frac{\omega_i}{\omega_j}$ is a rational number for any $i, j \in \{1, 2, \cdots, m\}$. Also the following inequalities need to be satisfied: $\omega_i \neq \omega_j$ and $\omega_i + \omega_j \neq \omega_k$ for distinct $i, j, k \in \{1, 2, \cdots, m\}$.

For the convenience of notation, it is denoted that $\omega_{min} = \min\{\omega_1, \cdots, \omega_m\}$. Following the similar steps in Section III, the close-loop system in Fig.2 after the same coordinate change becomes

$$\begin{aligned}\dot{\tilde{\mathbf{x}}} =&\mathbf{f}(\tilde{\mathbf{x}} + \mathbf{x}^*) + \mathbf{g}(\tilde{\mathbf{x}} + \mathbf{x}^*)F(\tilde{\boldsymbol{\theta}} + \boldsymbol{\theta}^* + a\mathbf{d}(\omega t))\\ \dot{\tilde{\boldsymbol{\theta}}} =&\frac{K}{a}(L_{\mathbf{f}}h + L_{\mathbf{g}}hF(\tilde{\boldsymbol{\theta}} + \boldsymbol{\theta}^* + a\mathbf{d}(\omega t))\mathbf{d}(\omega t),\end{aligned} \tag{11}$$

for all $\tilde{\mathbf{x}}(t_0) \in \mathcal{R}^n$, $\tilde{\boldsymbol{\theta}}(t_0) \in \mathcal{R}^m$, and $t \geq t_0 \geq 0$.

Let $\varepsilon = \frac{K}{a\omega_{min}}$. The following corollary summarizes the stability properties for the system (11).

Corollary 1: Suppose the Assumptions 1, 3, 4 hold for the system (11), and the dither signal $\mathbf{d}(\omega t)$ satisfies Condition 1. Given a positive triplet (Δ, ν, K), there exist $\beta_1, \beta_2 \in \mathcal{KL}$ and a positive constant $a^* > 0$ such that for any $a \in (0, a^*)$, there exists $\varepsilon^* > 0$ such that any $\varepsilon \in (0, \varepsilon^*)$, the solutions of the system (11) satisfy:

$$\begin{aligned}|\tilde{\mathbf{x}}(t)| \leq& \beta_1(|\tilde{\mathbf{x}}(t_0)|, t - t_0) + \nu,\\ \left|\tilde{\boldsymbol{\theta}}(t)\right| \leq& \beta_2(\left|\tilde{\boldsymbol{\theta}}(t_0)\right|, K(t - t_0)) + \nu,\end{aligned} \tag{12}$$

for all $t \geq t_0 \geq 0$ and $\left\|\begin{bmatrix}\tilde{\mathbf{x}}(t_0)\\ \tilde{\boldsymbol{\theta}}(t_0)\end{bmatrix}\right\| \leq \Delta$.

The proof of Corollary 1 can directly follow the steps used in the proof of Theorem 1 (see Appendix) as well as the PE condition (see Condition 1 and discussions in [1, 6]). Thus it is omitted due to space limitation.

Remark 5: It is highlighted that the proposed fast ESC in Fig.1 and Fig.2 approximate the gradient of $F(\cdot)$. This can be extended to other optimization techniques based ESC, for example, the Newton-based method in [6], the accelerated gradient method in [16], and so on. Our future work will investigate different fast ESC using other off-the-shelf optimization algorithms.

V. Simulation Examples

This section presents two simulation examples to illustrate the performance of the fast ESC. The first example demonstrates the performance of the proposed fast ESC with different tuning parameters for a SISO generalized Hammerstein system, while the second example illustrates how to choose appropriate dither signals, which satisfy Condition 1, for a cost function with multiple inputs.

A. A single-input case

The following single-input generalized Hammerstein model is considered:

$$\begin{aligned}u =&F(\theta)\\ =&\begin{cases}e^{-0.5(\theta-1)^2}, & |\theta - 1| \leq c,\\ -e^{-c^2}(0.5(\theta-1)^2 - 0.5c^2 - 1), & |\theta - 1| > c,\end{cases}\\ \dot{x} =&-x + u,\\ y =&x - 0.5\arctan(x),\end{aligned} \tag{13}$$

where c is a large positive constant. The initial state of (13) is chosen as $x(0) = -1$, and the initial state of the ESC is $\hat{\theta}(0) = -1$. The cost function $F(\cdot)$ is continuous differentiable and satisfies Assumption 3 with a unique optimum $\theta^* = 1$. It is also verified that Assumption 2 holds for this system with $x^* = 1$. Moreover, it is checked that $0.5 \leq L_g h = \frac{x^2+0.5}{x^2+1}$, satisfying Assumption 1. Thus, the proposed fast ESC scheme in Fig.1 is applicable.

We first check the convergence property of the proposed fast ESC in Fig.1. It is highlighted that the convergence speed of $\hat{\theta}$ only depends on the choice of the parameter K, as stated in Theorem 1. We fix $a = 0.1$, $\varepsilon = \frac{K}{a\omega} = 0.1$, and choose $K = 8$, 10, 20, 50, 100, 200 respectively. As shown in Fig.3, a larger K leads to a faster convergence speed for $\hat{\theta}$, which is consistent with the results in Theorem 1 while the convergence of $\mathbf{x}$ will be slightly improved with larger K (mainly depends on the dynamics of the system).

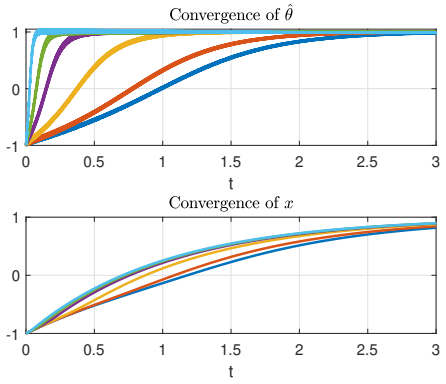

Fig. 3. The trajectories of $\hat{\theta}$ and x by using the proposed fast ESC with different value of K. The value of K from right to left: 8, 10, 20, 50, 100, 200.

Next we use the classic perturbation-based ESC [19] for the system (13) for comparison. By selecting $a = b = 0.2$, $k = 0.2$, $\omega = 0.5$ (see [19]), the trajectories of $\hat{\theta}$ and x are shown in Fig.4. We can see that the classic ESC has slow convergence speed from Fig.4 due to necessary time-scale separation.

Last we will discuss how to improve the robustness with respect to measurement noises. As mentioned in Remark 2, usually the non-causal differentiator is used along

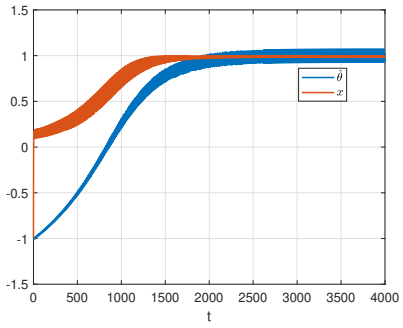

Fig. 4. The trajectories of $\hat{\theta}$ and x from the classic ESC [19].

with a proper filter. A high frequency disturbance $d_{out} = 0.001\sin(1800t)$ is added to the output of the plant. We choose a set of parameters $a = 0.1$, $\delta = \frac{K}{a\omega} = 0.1$, and $K = 10$. Here $\omega = 1000rad/s$. In order to reduce the influence of this noise to the fast ESC, a simple low-pass filter: $\frac{1}{\frac{s}{1200}+1}$ is connected with the differentiator. Fig.5 shows the trajectory of $\hat{\theta}$ with/without a low-pass filter, indicating the performance improvement in the presence of high frequency noises when a low-pass filter is used.

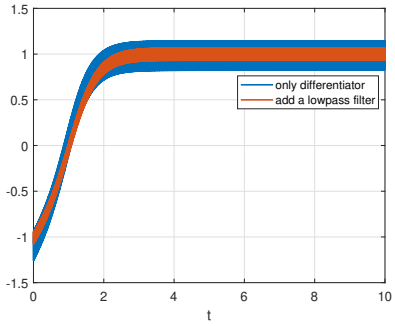

Fig. 5. The trajectories of $\hat{\theta}$ by using the proposed fast ESC with a low-pass filter.

B. A multi-input case

Now we consider a generalized Hammerstein model with multiple inputs:

$$\begin{aligned} u =& -0.1(\theta_1 - 1)^2 + -0.2(\theta_2 - 2)^2, \\ \dot{x} =& -x + u, \\ y =& x - 0.5\arctan(x). \end{aligned} \tag{14}$$

We choose the initial state of (14) as $x(0) = -1$, and the initial state of the ESC is chosen as $\hat{\theta}_1(0) = -1$, $\hat{\theta}_2(0) = 0$. The global maximum of the nonlinear mapping in (14) is $\theta_1^* = 1$, $\theta_2^* = 2$, and the equilibrium of the nonlinear system $x^* = 1$. Assumptions $1-3$ are verified for the system (14).

In the simulation, tuning parameters are selected as: $a = 0.1$, $\varepsilon = \frac{K}{a\omega_{min}} = 0.1$, $K = 20$. The frequencies of the dithers are $\omega_1 = \omega_{min}$ and $\omega_2 = 1.05\omega_{min}$, which satisfies Condition 1. Fig.6 shows the trajectories of $\hat{\theta}_1$, $\hat{\theta}_2$ and x. From Fig.6, we can see that there is no time-scale separation between input $\hat{\boldsymbol{\theta}}$ and state x, which is consistent results with respect to Corollary 1.

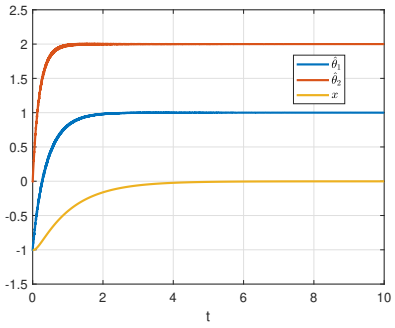

Fig. 6. The trajectories of $\hat{\theta}_1$, $\hat{\theta}_2$ and x for multi-input system (14) by using the proposed fast ESC

VI. Conclusion

This work proposed the fast ESC for a class of generalized Hammerstein systems. The proposed fast ESC can seek the optimum (either a scalar or a vector) arbitrarily close from any compact set in which the initial state stays. This can be achieved by appropriate tuning of parameters of the proposed fast ESC with the knowledge of the relative degree of the nonlinear dynamics. As shown in this work, the convergence speed of the proposed ESC approach can be significantly improved compared to classic perturbation-based ESC, in which time-scale separation between the dynamics and the updating law is needed.

Appendix

Proof of Theorem 1.

Averaging techniques [11, Chapter 10] will be used to analyze the stability properties of the close-loop system (6). Let $\tau = \omega t$. In the new time "τ", the system (6) becomes

$$\begin{aligned} \frac{d\tilde{\mathbf{x}}}{d\tau} =& \frac{1}{\omega}\left(\mathbf{f}(\tilde{\mathbf{x}} + \mathbf{x}^*) + \mathbf{g}(\tilde{\mathbf{x}} + \mathbf{x}^*)F(\tilde{\theta} + \theta^* + a\sin(\tau))\right), \\ \frac{d\tilde{\theta}}{d\tau} =& \frac{K}{a\omega}(L_{\mathbf{f}}h + L_{\mathbf{g}}hF(\tilde{\theta} + \theta^* + a\sin(\tau)))\sin(\tau). \end{aligned} \tag{15}$$

For any given K, by selecting a satisfying $\frac{1}{\omega} \leq \frac{K}{a\omega}$ and $0 < a \ll 1$, we can define $\varepsilon = \frac{K}{a\omega}$ such that (15) has a clear time-scale separation between the fast time-varying part $\sin(\tau)$ and slower dynamics in both $\tilde{\mathbf{x}}$ and $\tilde{\theta}$. Thus the standard averaging techniques in [11, Chapter 10] is directly applicable with $\varepsilon = \frac{K}{a\omega}$. For simplicity of notation, it is defined $c_0 = \frac{a}{K} < 1$, $\eta_1 := \mathbf{f}(\tilde{\mathbf{x}} + \mathbf{x}^*)$, $\eta_2 := \mathbf{g}(\tilde{\mathbf{x}} + \mathbf{x}^*)$, $\eta_3 := F(\tilde{\theta} + \theta^*)$, $\eta_4 := \frac{dF}{d\theta}(\tilde{\theta} + \theta^*)$, $\eta_5 := \frac{d^2F}{d\theta^2}(\tilde{\theta} + \theta^*)$.

By using Taylor series expansion, the system (15) becomes

$$\begin{aligned}\frac{d\tilde{\mathbf{x}}}{d\tau} &= \varepsilon c_0\left(\eta_1+\eta_2\cdot\eta_3+a\eta_2\cdot\eta_4\sin(\tau)+O\left(a^2\right)\right)\\ \frac{d\tilde{\theta}}{d\tau} &= \varepsilon\left(L_{\mathbf{f}}h\cdot\sin(\tau)+L_{\mathbf{g}}h\cdot\eta_3\sin(\tau)+a\eta_4\sin^2(\tau)\right)\\ &\quad+\varepsilon\left(0.5a^2\eta_5\sin^3(\tau)+O\left(a^3\right)\right),\end{aligned}\tag{16}$$

with its following averaged system:

$$\begin{aligned}\frac{d\mathbf{x}_{av}}{d\tau} =&\varepsilon c_0\left(\mathbf{f}(\mathbf{x}_{av}+\mathbf{x}^*)+\mathbf{g}(\mathbf{x}_{av}+\mathbf{x}^*)F(\theta_{av}+\theta^*)+O\left(a^2\right)\right)\\ \frac{d\theta_{av}}{d\tau} =&\varepsilon\left(0.5aL_{\mathbf{g}}h\frac{dF}{d\theta_{av}}(\theta_{av}+\theta^*)+O\left(a^3\right)\right).\end{aligned}\tag{17}$$

The averaged system (17) in time-scale t has the following form:

$$\begin{aligned}\dot{\mathbf{x}}_{av} =&\mathbf{f}(\mathbf{x}_{av}+\mathbf{x}^*)+\mathbf{g}(\mathbf{x}_{av}+\mathbf{x}^*)F(\theta_{av}+\theta^*)+O\left(a^2\right)\\ \dot{\theta}_{av} =&0.5KL_{\mathbf{g}}h\frac{dF}{d\theta_{av}}(\theta_{av}+\theta^*)+O\left(a^2\right).\end{aligned}\tag{18}$$

For a given compact set Δ, for any $\mathbf{x}\in\Delta$, there exist two constants $M_\delta>0$, $M_\Delta>0$ such that

$$M_\delta\le L_{\mathbf{g}}h\le M_\Delta,\tag{19}$$

from the continuity of $L_{\mathbf{g}}h$ and Assumption 1. By using Assumption 3, following the similar steps used in [18], we can conclude that θ_{av} dynamics in (18) is semi-globally practically asymptotically stable (SPA) in a (see [19, Definition 1] for the detailed definition of SPA in parameter a). By using Assumption 2, once θ_{av} is convergent, we can conclude the stability properties of $\mathbf{x}_{av}$. That is for a positive pair (Δ,ν), there exists $a^*>0$ such that for any $a\in(0,a^*)$, the convergence properties of the averaged system (17) can be captured by two class $\mathcal{KL}$ functions.

For the fixed $a\in(0,a^*)$, by applying the averaging result [18, Lemma 1], there exists $\varepsilon^*>0$ such that inequalities (7) hold for any $\varepsilon\in(0,\varepsilon^*)$. And this completes the proof. □

REFERENCES

[1] K. B. Ariyur and M. Krstic. "Analysis and design of multivariable extremum seeking". In: *Proceedings of the 2002 American Control Conference*. Vol. 4. 2002, pp. 2903–2908.

[2] W. Bamberger and R. Isermann. "Adaptive on-line steady-state optimization of slow dynamic processes". In: *Automatica* 14.3 (1978), pp. 223–230.

[3] P. Binetti et al. "Control of formation flight via extremum seeking". In: *Proceedings of the 2002 American Control Conference*. Vol. 4. 2002, pp. 2848–2853.

[4] Q. Chen et al. "Decentralized PID control design for magnetic levitation systems using extremum seeking". In: *IEEE Access* 6 (2018), pp. 3059–3067.

[5] C. S. Draper and Y. T. Li. *Principles of optimalizing control systems and an application to the internal combustion engine*. American Society of Mechanical Engineers, 1951.

[6] A. Ghaffari, M. Krstić, and D. Nešić. "Multivariable Newton-based extremum seeking". In: *Automatica* 48.8 (2012), pp. 1759–1767.

[7] M. Guay. "A perturbation-based proportional integral extremum-seeking control approach". In: *IEEE Transactions on Automatic Control* 61.11 (2016), pp. 3370–3381.

[8] M. Guay and K. T. Atta. "Dual mode extremum-seeking control via Lie-bracket averaging approximations". In: *2018 Annual American Control Conference (ACC)*. 2018, pp. 2972–2977.

[9] C. I. Hoarcă and F. M. Enescu. "On the energy efficiency of standalone fuel cell/renewable hybrid power sources Part I: Simulation results for constant load profile without RES power". In: *2018 International Conference on Applied and Theoretical Electricity (ICATE)*. 2018, pp. 1–6.

[10] A. Isidori. *Nonlinear Control Systems*. Springer Berlin Heidelberg, 1995.

[11] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.

[12] S. Z. Khong et al. "Extremum seeking of dynamical systems via gradient descent and stochastic approximation methods". In: *Automatica* 56 (2015), pp. 44–52.

[13] M. Krstić and H. H. Wang. "Stability of extremum seeking feedback for general nonlinear dynamic systems". In: *Automatica* 36.4 (2000), pp. 595–601.

[14] W. H. Moase and C. Manzie. "Fast extremum-seeking for Wiener–Hammerstein plants". In: *Automatica* 48.10 (2012), pp. 2433–2443.

[15] W. H. Moase and C. Manzie. "Semi-global stability analysis of observer-based extremum-seeking for Hammerstein plants". In: *IEEE Transactions on Automatic Control*. 2012.

[16] J. I. Poveda and N. Li. "Robust hybrid zero-order optimization algorithms with acceleration via averaging in time". In: *Automatica* 123 (2021), p. 109361.

[17] R. C. Shekhar, W. H. Moase, and C. Manzie. "Discrete-time extremum-seeking for Wiener–Hammerstein plants". In: *Automatica* 50.12 (2014), pp. 2998–3008.

[18] Y. Tan, D. Nešić, and I. Mareels. "On non-local stability properties of extremum seeking control". In: *IFAC Proceedings Volumes* 38.1 (2005), pp. 550–555.

[19] Y. Tan, D. Nešić, and I. Mareels. "On non-local stability properties of extremum seeking control". In: *Automatica* 42.6 (2006), pp. 889–903.

[20] Y. H. Zhang et al. "Robots looking for interesting things: Extremum seeking control on saliency maps". In: *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2011, pp. 1180–1186.