Author/s:
Tan, Y;Nesic, D;Mareels, I;Astolfi, A

Title:
On global extremum seeking in the presence of local extrema

Date:
2006-01-01

Citation:
Tan, Y., Nesic, D., Mareels, I. & Astolfi, A. (2006). On global extremum seeking in the presence of local extrema. PROCEEDINGS OF THE 45TH IEEE CONFERENCE ON DECISION AND CONTROL, VOLS 1-14, pp.5667-+. IEEE. https://doi.org/10.1109/cdc.2006.377791.

Persistent Link:
https://hdl.handle.net/11343/299660

On global extremum seeking in the presence of local extrema

Ying Tan, Dragan Nešić, Iven Mareels and Alessandro Astolfi

***Abstract*— We analyze global extremum seeking in the presence of local extrema for static nonlinear maps controlled by a scalar extremum seeking scheme that was recently proposed in [14]. Sufficient conditions are given under which it is possible to tune the controller parameters to achieve convergence to an arbitrarily small neighbourhood of the global extremum from an arbitrarily large set of initial conditions. These sufficient conditions are shown to hold always for 4^{th} order polynomials with two maxima but a 6^{th} order polynomial is presented that invalidates these conditions. On the other hand, extensive computations show that most 6^{th} order polynomials and many other functions satisfy all our conditions. Several examples provide insights and highlight the potential difficulties that one would face when trying to generalize our results.**

I. INTRODUCTION

The main goal in extremum seeking (ES) control is to find an extremum value of an unknown nonlinear mapping. This is an old method but the first rigorous local stability analysis for a class of ES schemes was provided recently in [1] and later extended to semi-global stability analysis in [15]. Stability of a different class of ES controllers was recently presented in [12]. There has been a renewed interest in this research area [4], [5], [11], [12], [13] that lead to numerous practical implementations of the scheme, as well as its better theoretical understanding.

While it has been often demonstrated that ES controllers work well in simulations, experiment or real applications, a full understanding of their convergence properties and robustness is still lacking. Global extremum seeking in absence of local extrema for a class of extremum schemes was rigorously analyzed in [14], [15]. On the other hand, it was often observed by the users of extremum seeking controllers that by tuning the amplitude of the excitation (dither) signal it is possible to pass through a local extremum and converge to the global one. In other words, global extremum seeking is possible in the presence of local extrema in certain situations. However, rigorous analysis of this problem appears to be lacking in the literature.

It is the purpose of this paper to present the first analysis of global extremum seeking in the presence of local extrema for the simplest possible situation: we consider a static scalar map controlled by the scalar ES controller proposed in [14]; moreover, as the amplitude of the excitation signal is changed on-line, the overall closed loop system is second order. This is the simplest possible variant of ES schemes considered in [1] that is amenable to rigorous analysis. Nevertheless, the problem is quite hard even in this simplest case and we illustrate this by several examples. Moreover, we present a set of sufficient conditions under which the proposed ES scheme yields global extremum seeking despite local extrema. Our conditions are shown to hold always for 4^{th} order polynomials, whereas we present and example of a 6^{th} order polynomial that does not satisfy these conditions. On the other hand, extensive computations show that most 6^{th} order polynomials, many higher order polynomials and more general functions satisfy all our conditions. Furthermore, when our conditions hold, our main result outlines a tuning strategy for the ES controller that yields convergence to an arbitrarily small neighbourhood of the global extremum from an arbitrarily large set of initial conditions. Hence, we believe that our results will be useful to the users of extremum seeking control and, moreover, they may motivate further research into this challenging area.

The paper is organized as follows. Blah.

II. PRELIMINARIES AND PROBLEM FORMULATION

The set of real numbers is denoted as R and the set of integers is denoted as N. The continuous function $\beta : R_{\geq 0} \times R_{\geq 0} \to R_{\geq 0}$ is of class $\mathcal{KL}$ if it is nondecreasing in its first argument and strictly decreasing to zero in its second argument. For a nonlinear smooth function $h : R^n \to R$ we denote $D_i^j h := \frac{\partial^j h}{\partial x_i^j}$ where $i \in \{1, 2, \ldots, n\}$ and $j \in N$. When $j = 1$ or $i = 1$, we often omit this argument, e.g. we write $D_i h := D_i^1 h$.

In this paper, we consider the scalar extremum seeking scheme shown in Figure 1 that was first introduced in [14]. The excitation signal $a\sin(t)$ is added to nonlinear mapping to get probing while the multiplication (modulation)

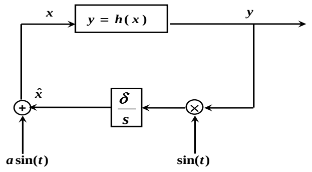

Fig. 1. The extremum seeking feedback scheme from [14]

This research was supported by the Australia Research Council under the Discovery Project DP0344784

Y. Tan, D. Nešić and I.M.Y. Mareels are with The Department of Electrical and Electronics, The University of Melbourne, Parkville, VIC 3010 `y.tan@ee.unimelb.edu.au`

Prof. Astolfi is with Department of Electrical and Electronic Engineering and Centre for Process Systems Engineering, Imperial College, Exhibition Road London SW7-2BT, United Kingdom

of output and the excitation signal ($sin(t)$) extracts the gradient of the unknown mapping $h(\cdot)$. The dynamics of the ES system in Figure 1 can be written as

$$\dot{x} = \delta h(x + a\sin(t))\sin(t) \overset{\triangle}{=} \delta f(t, x, a), \tag{1}$$

where $h(\cdot)$ is a smooth function. The state x is in R and $a > 0$ is the amplitude of the excitation signal. The nonlinear function $h(\cdot)$ satisfies the following assumption:

Assumption 1: There exists a unique global maximum $x^* \in R$ of $h(\cdot)$ such that[1]

$$Dh(x^*) = 0; \qquad D^2 h(x^*) < 0 \ , \tag{2}$$

$$h(x^*) > h(x) \qquad \forall x \in R,\ x \neq x^* \tag{3}$$

○

Note that (2) and (3) in Assumption 1 indicate that x^* is the global maximum of the nonlinear mapping $h(\cdot)$. Other than this global maximum, there may also exist local maxima that satisfy (2). This assumption is different from [14, Assumption 3], in which $x^* \in R$ is the unique maximum which satisfies (2). In this paper, we discuss sufficient conditions to ensure that the global maximum would be found if the amplitude of the excitation signal is tuned adaptively. To this end, we introduce the following law for adaptation of the amplitude of the excitation signal:

$$\dot{a} \ = \ -\delta \cdot \epsilon \cdot g(a), \quad a(0) = a_0 > 0, \tag{4}$$

where $g(\cdot)$ is a locally Lipschitz function that is zero at zero and positive otherwise and the strictly positive parameters ϵ, δ and a_0 are to be chosen by the designer. The simplest choice is $g(a) = a$. Note that the overall system (1), (4) is second order and its block diagram is given in Figure 2.

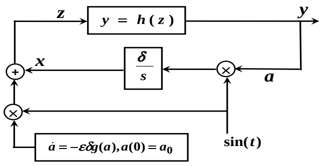

Fig. 2. The closed loop with the proposed extremum seeking controller

For the system (1), (4) we can write its "averaged" system by using:

$$f_{av}(x, a) := \frac{1}{2\pi}\int_0^{2\pi} f(t, x, a)dt \tag{5}$$

where f comes from (1). Indeed, using the above definition, we can analyze the closed loop system (1), (4) via the following auxiliary average system:

$$\begin{aligned} \dot{x} &= \delta f_{av}(x, a) \\ \dot{a} &= -\delta\epsilon g(a), \qquad a(0) = a_0 > 0 \ . \end{aligned} \tag{6}$$

Remark 1: By introducing the new time $\tau := \epsilon\delta t$, we can rewrite the above equations as follows:

$$\begin{aligned} \epsilon \cdot \frac{dx}{d\tau} &= f_{av}(x, a) \\ \frac{da}{d\tau} &= -g(a), \qquad a(0) = a_0 > 0 \ , \end{aligned} \tag{7}$$

which exhibits time scale separation and appears to be in standard singular perturbation form [8, Section 9.1]. However, there are three reasons why we do not use the standard singular perturbation techniques here. First, in our case the equation:

$$0 = f_{av}(x, a) \tag{8}$$

may not have k *isolated* roots $x = \ell_i(a)$, which is required of the standard form. Indeed, some of the roots may only be defined for $a \in [0, \bar{a}]$ and such that for some i and j we have $\ell_i(a) \neq \ell_j(a), a \in [0, \bar{a})$ and $\ell_i(\bar{a}) = \ell_j(\bar{a})$ (see Example ??). Moreover, it will be shown later that there exists a continuous function $p(x, a)$ such that we can write:

$$f_{av}(x, a) = a \cdot p(x, a) \tag{9}$$

and this means that we will be unable to prove stability of the boundary layer system *uniform in* a that is a standard assumption in the singular perturbation literature. Finally, we are interested in convergence properties of this system that is initialized from a set of initial conditions satisfying $a(0) = a_0$ which is a weaker property from the standard stability properties considered in the singular perturbation literature. Hence, we will state and prove results directly without appealing to the rich literature on singular perturbations. ○

Before we state our main results, we state another assumption that characterizes solutions of the equation (8).

Assumption 2: There exists an isolated root $x = \ell(a)$ of the equation (8) with the following properties:

1) ℓ is defined for all $a \geq 0$, it is continuous and $D_1 p(\ell(a), a) < 0$, $\forall a \geq 0$, where $p(x, a)$ is defined in (9).
2) There exists $a^* > 0$, such that for all $a \geq a^*$, $x = \ell(a)$ is the unique root of (8).
3) $\ell(0) = x^*$, where x^* comes from Assumption 1.

○

Remark 2: Note that if $h(\cdot)$ was a known mapping, it would be very easy to check conditions of Assumption 2. Indeed, one needs to plot the bifurcation diagram using (8) that shows how equilibria of the x subsystem of the average system (6) change as we vary the parameter a and verify the conditions by inspecting plot. However, since the standing assumption in extremum seeking control is that $h(\cdot)$ is *unknown*, then having results that guarantee conditions of Assumption 2 for *classes of functions* $h(\cdot)$ is more useful to the users of extremum seeking control (as in this case we do not need to plot the bifurcation diagram). Indeed, it may be known that $h(\cdot)$ belongs to a certain class of functions but its exact description may be unknown. Conditions

---

[1] In this paper we assume that $h(\cdot)$ possesses a global maximum but we can deal with functions $h_1(\cdot)$ with a global minimum by defining $h(\cdot) := -h_1(\cdot)$ and applying our results unchanged.

in Assumption 2 are hard to check in general for large classes of functions. We show that when $h(\cdot)$ is any 4^{th}-order polynomial with two maxima, these conditions are satisfied automatically. Moreover, we present an example of a 6^{th} order polynomial that invalidates some of these conditions. We have carried out extensive computations (plotted the bifurcation diagrams) and found that many higher order polynomials, as well as general functions satisfy all conditions of Assumption 2. Note that our results will apply to any system for which Assumptions 1 and 2 hold and, hence, they are general.

III. MAIN RESULT

In this section we present our main result (Theorem 1) that states that the closed loop system (1), (4) achieves semi-global practical convergence to the global extremum if Assumptions 1 and 2 are satisfied. Hence, semi-global practical extremum seeking is achieved despite the presence of local extrema (cf. Assumption 1). Moreover, we show that general 4^{th} order polynomials that satisfy Assumption 1 satisfy also all conditions in Assumption 2. On the other hand, we present an example of a 6^{th} order polynomial that does not satisfy some of the conditions in Assumption 2. Throughout this section we present many examples that illustrate and highlight the important issues and provide intuition that the users of extremum seeking control will find useful. Our main result is stated next.

Theorem 1: Suppose that Assumptions 1 and 2 hold. Then, for any strictly positive (Δ, ν) and $a_0 > a^*$ there exist $\beta = \beta_{a_0,\Delta,\nu} \in \mathcal{KL}$ and $\epsilon^* = \epsilon^*(a_0, \Delta, \nu) > 0$ and for any $\epsilon \in (0, \epsilon^*)$ there exists $\delta^* = \delta^*(\epsilon) > 0$ such that for any such a_0, ϵ and $\delta \in (0, \delta^*)$ we have that for all $(x(t_0), a(t_0))$ satisfying $a(t_0) = a_0$ and $|x(t_0) - \ell(a_0)| \leq \Delta$ and all $t \geq t_0 \geq 0$ the solutions of the system (1), (4) satisfy:

$$|x(t) - \ell(a(t))| \leq \beta(|x(t_0) - \ell(a(t_0))|, \delta(t - t_0)) + \nu \ . \tag{10}$$

◦

Remark 3: A consequence of Theorem 1 is that we can tune the extremum seeking controller to achieve:

$$\limsup_{t\to\infty} |x(t) - \ell(a(t))| \leq \nu$$

from an arbitrarily large set of initial conditions and for arbitrarily small $\nu > 0$. Moreover, from (4) it is obvious that there exists $\beta_a \in \mathcal{KL}$ with $\beta_a(s, 0) = s$, such that for all $a(t_0) = a_0 \in R$ we have:

$$|a(t)| \leq \beta_a(a(t_0), \epsilon \cdot \delta \cdot (t - t_0)), \ \forall t \geq t_0 \geq 0 \ , \tag{11}$$

and since $\ell(\cdot)$ is continuous and $\ell(0) = x^*$, this implies

$$\lim_{t\to\infty} \ell(a(t)) = x^* \ .$$

Hence, we can conclude that

$$\limsup_{t\to\infty} |x(t) - x^*| \leq \nu \ ,$$

which implies semi-global practical extremum seeking since x^* is the global extremum of $h(\cdot)$. We believe that this is the first rigorous result of this kind in the literature. ◦

Remark 4: The averaged model (7) suggest a two time scale dynamics when ϵ is very small. This is indeed the case, as can be seen from (10) and (11). Indeed, the solutions first converge with the rate proportional to δ to a small neighbourhood of the set $\mathcal{L} := \{(x, a) : x - \ell(a) = 0\}$ (fast transient given by (10)) and then with the speed proportional to $\epsilon \cdot \delta$ to a neighbourhood of the point $(x, a) = (x^*, 0)$ (slow transient given by (11)). Moreover, during the slow transient, the solutions stay in the ν-neighbourhood of the set $\mathcal{L}$. ◦

Remark 5: The controller in Theorem 1 is more general than the controller used in the main results of [14] where local extrema were not allowed in the analysis. The main difference between the two controllers is that the amplitude of the excitation signal in Theorem 1 is time varying, whereas in main results of [14] the amplitude is *fixed*. ◦

Remark 6: Theorem 1 presents a tuning mechanism for the controller parameters (choice of ϵ, δ) and its initialization (choice of a_0) that guarantees semi-global practical convergence to the global extremum despite the presence of local extrema. Simulations in our examples illustrate that such convergence is indeed achieved.

We note that since the static mapping $h(\cdot)$ is not known, it is in general not possible to check *a priori* whether Assumption 2 holds, let alone analytically compute the values of a_0, ϵ^* and δ^*. However, our result suggests that if there is some evidence that Assumption 2 may hold, then increasing sufficiently a_0 and reducing sufficiently ϵ and δ will indeed result in global convergence. In practice, determining how large a_0, ϵ, δ should be, may have to be determined through experimenting. ◦

Remark 7: Note that from the robustness point of view, it is not desirable to reduce the amplitude of the excitation signal to zero since small perturbations may force the solutions to diverge far from the global extremum. Our results can be restated for the case when we have that $\lim_{t\to\infty} a(t) = \underline{a} > 0$ but they are omitted for reasons of brevity. ◦

Remark 8: We note that Assumption 2 imposes conditions on the bifurcation diagram that is defined by the equation (8). The bifurcation diagram is a real algebraic variety in the case when f_{av} is a polynomial or a more general set for general functions f_{av}. The idea behind this condition is to ensure that for large amplitudes a the average map has a unique equilibrium that is a global extremum, given by $x = \ell(a)$. Moreover, in order for our proposed control strategy to be successful we require this branch to be continuous and connected to the global maximum, i.e. $\ell(0) = x^*$. While this always holds for 4^{th} order polynomials (this is shown next), we present an example of a 6^{th} order polynomial where this does not hold (Example 3). Moreover, after extensive plotting of bifurcation diagrams we observed that for 6^{th} order polynomials all conditions in Assumption 2 hold most of the time (i.e. counterexamples are hard to construct). Similarly, these conditions also hold for many higher order polynomials or general functions $h(\cdot)$, which indicates that our results are quite general. ○

Next, we present an example (4^{th} order polynomial) that satisfies all conditions of Assumptions 1 and 2 and to which our main result in Theorem 1 applies.

Example 1: Consider $h(x) = -x^4 + \frac{8}{15}x^3 + \frac{5}{6}x^2 + 10$ that has a global maximum at $x = 1$ and a local maximum at $x = -0.6$ as seen in Figure 3. Hence, all conditions in Assumption 1 hold. The average system is:

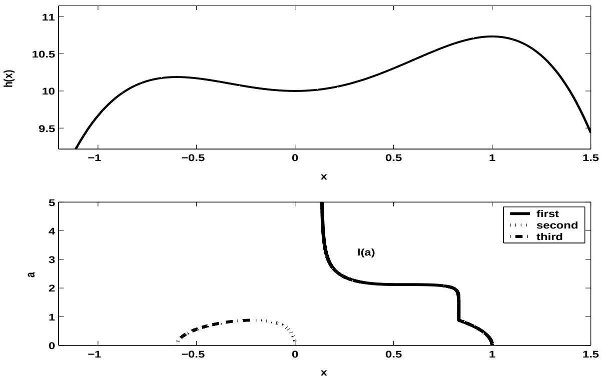

Fig. 3. Nonlinear function $h(x)$ and the corresponding zero plot of $f_{av}(x,a) = 0$

$$f_{av}(x,a) = \frac{a}{2}\left[-4x^3 + \frac{24}{15}x^2 + \frac{5}{3}x + a\left(-24x + \frac{48}{15}\right)\right].$$

Note that there exists a continuous root $\ell(a)$ of (8) approaching 1 when $a \to 0$. Moreover, it is not hard to check that this root satisfies all conditions of Assumption 2 and, hence, Theorem 1 applies. Indeed, consider the initial condition $x_0 = -1$ which is such that the local maximum $x = -0.6$ lays between the initial condition and the global maximum. Nevertheless, by choosing $a_0 = 3$, $\delta = 0.001$ and $\epsilon = 0.1$, the output of the system converges to a small neighborhood of the global maximum $x = 1$ as seen in Figure 4.

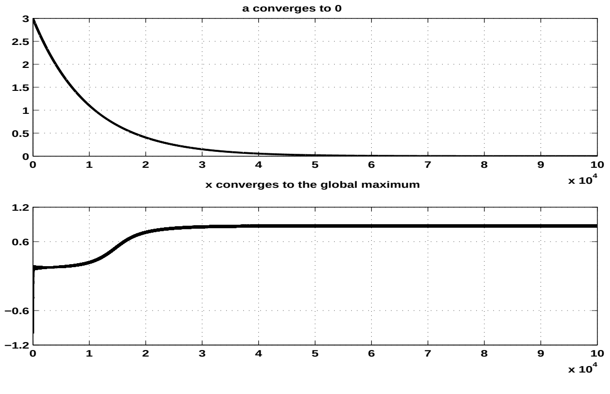

Fig. 4. The performance of the proposed ES scheme

○

Next, we show that arbitrary 4^{th}-order polynomials satisfying Assumption 1 also satisfy conditions in Assumption 2:

$$h(x) = \alpha_0 x^4 + \alpha_1 x^3 + \alpha_2 x^2 + \alpha_3 x + \alpha_4, \quad \alpha_0 = 1 \quad (12)$$

Proposition 1: Consider $h(\cdot)$ in (12). Suppose that Assumption 1 holds and that there is one local and one global maximum. Then, all conditions in Assumption 2 hold.

Note that since $\alpha_0 = -1$ in (12), we have that (2) in Assumption 1 always holds. However, this does not guarantee that (3) holds and we still need to assume this. Actually, if (3) in Assumption 1 does not hold while (2) holds, then we can not prove in general that Assumption 2 holds, as the following example illustrates.

Example 2: Consider the polynomial $h(\cdot) = -???$ that is such that there exist two points $x_1^* = x_2^*$ such that $h(x_1^*) = h(x_2^*)$ and $h(x) < h(x_1^*)$ for all $x \neq x_1^*$ and $x \neq x_2^*$. The plot of this function and its (pitchfork) bifurcation diagram are given in Figure 5. It is easy to see that there does not exist an *isolated* root of (8) that satisfies Assumption 2.

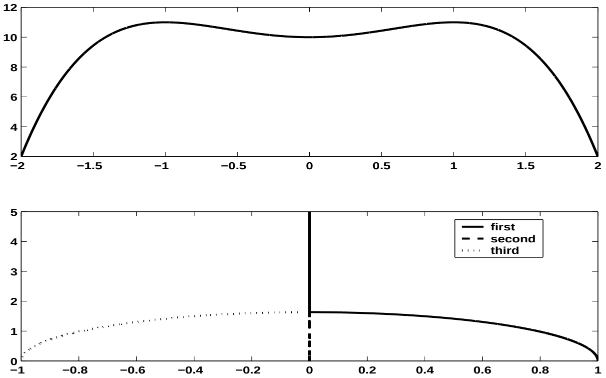

Fig. 5. Bifurcation diagram of a function that does not satisfy condition (3).

The bifurcation diagram of the 6^{th}-order polynomials are more complicated than those of 4^{th}-order polynomials. We have observed in simulations that for most of the cases (more than 80%), Assumption 2 holds. However, this is not true in general as the following example shows.

Example 3: For the following 6^{th}-order polynomial (see Figure 6):

$$\begin{aligned} h(x) &= -x^6 + \frac{1}{10}x^5 + \frac{623}{400}x^4 - \frac{659}{4000}x^3 \\ &\quad -\frac{11287}{20000}x^2 - \frac{259}{4000}x + \frac{637}{20000}. \qquad (13) \end{aligned}$$

which has maxima at -0.8985, $.5$, 0.8951. The global maximum occurs at $x^* = -0.8985$. The function $f_{av}(x,a)$ is:

$$a\left[\frac{1}{2}Dh(x) + \frac{a^2}{16}D^3h(x) + \frac{a^4}{384}D^5h(x)\right] \qquad (14)$$

whose bifurcation diagram is shown in Figure 6. It is clear that item 3 of Assumption 2 does not hold since $\ell(0) =$

0.8951 which is not a global maximum. This implies that our extremum seeking controller if tuned like in Theorem 1 would yield semi-global practical convergence of the *local maximum*.

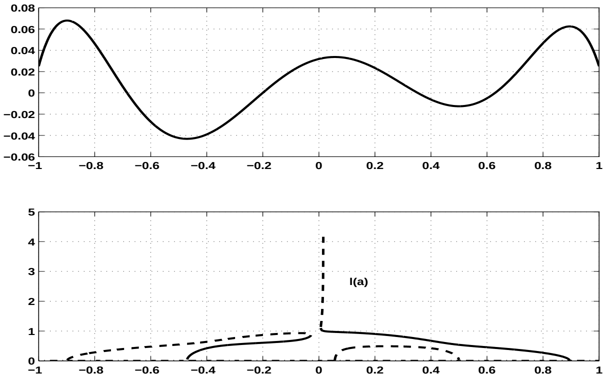

Fig. 6. The 6^{th} order polynomial and its bifurcation diagram for which Assumption 2 does not hold (sine wave dither).

Remark 9: It is not crucial that a sinusoidal signal is used as excitation (dither) in the extremum seeking controller. Indeed, one may use different dither signals, such as a square-wave or a sawtooth signals. Interestingly, changing the dither leads to a different average system f_{av} and the bifurcation diagram defined by (8) changes. An interesting consequence of this fact is that Assumption 2 may not be satisfied for one dither signal whereas it may be satisfied for a different dither. In our next example we revisit the system in Example 3 that did not satisfy Assumption 2 with a sinusoidal dither and show that the same system with a square wave dither satisfies Assumption 2. More analysis of how the choice of dither affects the convergence properties of extremum seeking controllers is given in [**?**].

Example 4: Consider again the function $h(\cdot)$ in Example 3 (see the top plot in Figure 6). Suppose that instead of the sine wave, we use a square wave dither in our controller, which is defined as follows

$$sq(t) := \begin{cases} -1 & t \in [kT, kT + T/2) \\ 1 & t \in [kT + T/2, kT + T) \end{cases} ,$$

where $k = 0, 1, \ldots$ and $T > 0$. Direct calculations yield the following $f_{av}(x, a)$:

$$a[Dh(x) + \frac{a^2}{6} D^3 h(x) + \frac{a^4}{120} D^5 h(x)], \tag{15}$$

which is different from (14). The bifurcation diagram is shown in Figure 7. Assumption 2 holds since $\ell(a)$ satisfies $\ell(0) = x^* = -0.8985$. Hence, if we use our controller and the tuning strategy from Theorem 1 we will obtain semi-global practical convergence to the *global maximum*.

Next we present an example of a more general function that also does not satisfy conditions of Assumption 2. In particular, item 3 in Assumption 2 does not hold.

We can exclude this example.

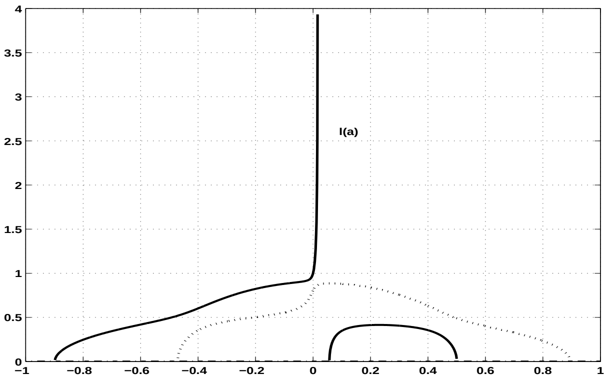

Fig. 7. The bifurcation diagram for $h(\cdot)$ from Example 3 for which Assumption 2 holds (square wave dither).

Example 5: Consider the function $h(x) = e^{\frac{1}{1+.02x.^2}} + e^{\frac{1}{1+5(x-15)^2}}$. This function has a global maximum at $x = 15$ and a local maximum at $x = 0$ (see Figure 8). The average system in this case is computed numerically and the bifurcation diagram plotted directly. There exists a continuous root $\ell(a)$ that is unique for large a whereas it approaches the local maximum $x = 0$, instead of the global one, when $a \to 0$. Hence, the item 3 of Assumption 2 does not hold. When the initial condition is close to the local maximum, no matter how we tune the controller parameters (δ, a_0, ϵ), the output of the proposed ES system converges to this local maximum. ◦

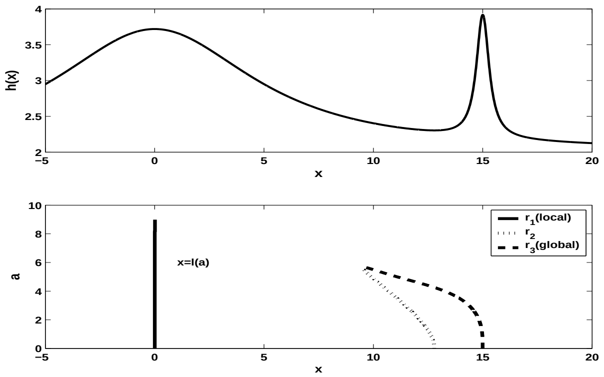

Fig. 8. Nonlinear function $h(x)$ and bifurcation diagram

IV. SUMMARY

We have presented an extremum seeking controller for static maps that achieves semi-global practical extremum seeking in presence of local extrema. Several examples were presented illustrating and highlighting various issues. We believe that these results will be of use to the control engineers that are using the extremum seeking control and may motivate further research into this challenging area.

REFERENCES

[1] K. B. Ariyur and M. Krstić, *Real-Time Optimization by Extremum-Seeking Control*, Wiley-Interscience, A John Wiley & Sons, Inc., Publication, 2003.

[2] L.Blum, F. Cucker, M. Shub and S.Smale *Complexity And Real Computation*. Springer-Verlag: New York, 1998.

[3] P.D. Christofides and A. R. Teel, “Singular perturbation and input-to-state stability” *IEEE Transactions on Automatic Control*, **41** pp. 1645–1650, 1996.

[4] M. Guaya and T. Zhang, ”Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties”, *Automatica* **39**, pp. 1283-1293, 2003.

[5] M. Guay, D. Dochain, M. Perrier, ”Adaptive extremum seekingcontrol of continuous stirred tank bioreactors with unknown growth kinetics”, *Automatica* **40**, pp. 881-888, 2004.

[6] M. Henruksen and J.R. Isbell, “On the continuity of the real roots of an algebraic equation”, *Proceddings of the American Mathematical Society* **4**, No. 3 pp 432-434, 1953.

[7] N.Jacobson, *Basic Algebra I*. W.H.Freeman and Co.: San Francisco, 1974.

[8] H. K. Khalil, *Nonlinear Systems*, Third edition Prentice Hall, Upper Saddle River, New Jersey, 2002.

[9] M. Marden, *Geometry of Polynomials*, American Mathematical Society Providence, Rhode Island, 1989.

[10] D. Nešić and A. R. Teel, “Input-to-state stability for nonlinear time-varying systems via averaging” *Math. Contr. Sign. Sys.*, **14** pp. 524–533, 2001.

[11] K. S. Peterson and A. G. Stefanopoulou, ”Extremum seeking control for soft landing of an electromechanical valve actuator”, *Automatica* **40**, pp. 1063-1069, 2004.

[12] D. Popović, M. Janković, S. Manger and A.R.Teel, ”Extremum seeking methods for optimization of variable cam timing engine operation”, *Proc. American Control Conference*, Denver, Colorado, June 4-6, pp. 3136–3141, 2003.

[13] A. R. Teel and D. I. Popović, “Solving smooth and nonsmooth extremum seeking problems by methods of nonlinear programming,” *Proc. of the American Control Conference*, June 2001, Washington, DC, 2001.

[14] Y. Tan; D. Nešić, I.M.Y. Mareels, “On non-local stability properties of extremum seeking control”, *To appear in Automatica*, 2006.

[15] Y. Tan; D. Nešić, I.M.Y. Mareels, “On non-local stability properties of extremum seeking control”, *Proc. 16th IFAC World Congress*, Prague, Czech Republic, pp. 2483–2489, 2005.

[16] A. R. Teel, L. Moreau and D. Nešić, “A unified framework for input-to-state stability in systems with two time scales”, IEEE Transactions on Automatic Control, **48**, No.9, pp. 1526–1544, 2003.

[17] A. R. Teel and D. Nešić, “Averaging with disturbances and closeness of solution” *Systems & Control Letters.*, **40** pp. 317–323, 2000.

[18] A. R. Teel, J. Peuteman, and D. Aeyels, “Semi-global practical asymptotic stability and averaging” *System & Control Letters*, **37** pp. 329–334, 1999.

V. APPENDIX

A. Proof of Theorem 1

First, we show that (9) holds. Using the Taylor Series Expansion, noting $\int_0^{2\pi} sin^{2i-1}(t)dt = 0$, $\forall i = 1,\ldots,$ we obtain that $f_{av}(x,a)$ is equal to:

$$a\left(\sum_{i=1}^{r} a^{2(i-1)} \cdot c_i \cdot D^{2i-1}h(x) + a^{2r} \cdot c_{r+1} \cdot R\right) =: a \cdot p(x,a), \quad (16)$$

where

$$c_i \triangleq \begin{cases} \frac{1}{(2i-1)!} \cdot \frac{1}{2\pi}\int_0^{2\pi} \sin^{2i}(t)dt & i = 1,\ldots,r \\ \frac{1}{(2r-1)!} & i = r+1 \end{cases}$$

$$R := \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \sin^{2r+1} t(1-s)^{(2r-1)} D^{2r}h(x + sa\sin t)dsdt$$

Before we prove the main result, we state several facts that follow directly from Assumption 2.

Proposition 2: Suppose that Assumption 2 holds. Then, for any $a_1 > 0$ there exists $\eta > 0$ and $\alpha_1 \in \mathcal{K}$ such that

$$\zeta \cdot p(\zeta + \ell(a), a) \le -\alpha_1(|\zeta|), \quad (17)$$

for all $|\zeta| \le \eta$ and $a \in [0, a_1]$.

Proof of Proposition 2: Introduce $\zeta := x - \ell(a)$ in (8) and note that the root $x = \ell(a)$ of (8) corresponds to $\zeta = 0$. Since $x = \ell(a)$ is an isolated root of (8), we have that for any $a_1 > 0$ there exists $\eta > 0$ such that for all $|\zeta| \le \eta$ and $a \in [0, a_1]$, $\zeta = 0$ is the only root of

$$p(\zeta + \ell(a), a) = 0$$

in this set. Moreover, since $p(\ell(a), a) = 0$, p is smooth and from the item 1 of the assumption we have $D_1 p(\ell(a), a) < 0$ for all $a \ge 0$, we conclude that

$$\zeta \cdot p(\zeta + \ell(a), a) < 0 \quad (18)$$

for all $|\zeta| \le \eta, \zeta \ne 0$ and $a \in [0, a_1]$. Since the set $|\zeta| \le \eta$ $a \in [0, a_1]$ is compact, we can define:

$$\kappa(s) := \min_{|\zeta| \in [s,\eta]} \min_{a \in [0,a_1]} [-\zeta \cdot p(\zeta + \ell(a), a)] \ ,$$

which is nondecreasing, continuous and $\kappa(0) = 0$ and, hence, we can find $\alpha_1 \in \mathcal{K}$ such that $\kappa(s) \ge \alpha_1(s)$ for all $s \in [0, \eta]$ that satisfies (17).

Proposition 3: Suppose that Assumption 2 holds. Then, for any $a_2 > a^*$ and any $\Delta > 0$ there exists $\alpha_2 \in \mathcal{K}$ such that

$$\zeta \cdot p(\zeta + \ell(a), a) \le -\alpha_2(|\zeta|), \quad (19)$$

for all $|\zeta| \le \Delta$ and $a \in [a^*, a_2]$.

Proof of Proposition 3: The proof of this proposition follows in a similar manner as the proof of Proposition 2, by noting that we also have that for arbitrary $a_2 > a^*$ and arbitrary Δ we have from the item 2 of the assumption that $\zeta = 0$ is a unique root of (18) on the set $a \in [a^*, a_2]$ and $|\zeta| \le \Delta$.

Proof of Theorem 1: First, we show that an appropriate bound can be obtained for the averaged system (6) if the (a_0, ϵ) are appropriately adjusted. Then, the conclusion of the theorem follows from recent results on averaging [17], [16], [18] and this part is omitted.

Consider the average system (6). We show for this system that for any strictly positive (D, d) and $a_0 > a^*$ there exist $\tilde{\beta} = \tilde{\beta}_{a_0,D,d} \in \mathcal{KL}$ and $\epsilon^* = \epsilon^*(a_0, D, d) > 0$ such that for any $\epsilon \in (0, \epsilon^*)$, all $(x(0), a(0))$ satisfying $a(0) = a_0$ and $|x(0) - \ell(a(0))| \le D$ and all $t \ge 0$ the solutions of the system (6) satisfy:

$$|x(t) - \ell(a(t))| \le \tilde{\beta}(|x(0) - \ell(a(0))|, \delta \cdot t) + d \ . \quad (20)$$

Let (D, d) and $a_0 > a^*$ be given and without loss of generality assume that $D \ge d$. Let $a_1 := a_0$ generate η and $\alpha_1 \in \mathcal{K}$ via Proposition 2. Let $\Delta := D$ and $a_2 := a_0$ generate $\alpha_2 \in \mathcal{K}$ via Proposition 3. Let $\alpha := \min\{\alpha_1, \alpha_2\}$. Let $c := \min\{\eta, \frac{1}{2}d\}$. To complete the proof, we introduce the following sets:

$$\mathcal{S}_1 := \{(\zeta, a) : |\zeta| \le D + d/2, a \in [a^*, a_0]\}$$

$$\mathcal{S}_2 := \{(\zeta, a) : |\zeta| \le c, a \in [0, a_0]\}$$

and we let $\mathcal{S} := \mathcal{S}_1 \cup \mathcal{S}_2$. These sets are shown in Figure 9.

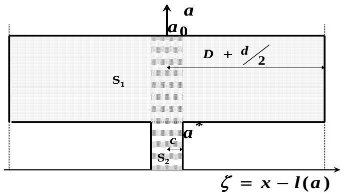

Fig. 9. Sets $\mathcal{S}_1$ and $\mathcal{S}_2$.

Let $\tilde{\beta}(s,t) \in \mathcal{KL}$ be the solution of the following differential equation:

$$\dot{\beta} = -\frac{a^*}{2}\alpha\left(\sqrt{2\beta}\right),\ \beta(0) = s\ ,$$

$\tilde{\beta}(s,t) := \sqrt{2\beta\left(\frac{s^2}{2},t\right)}$ and let $T > 0$ be such that

$$\tilde{\beta}\left(D + \frac{d}{2}, T\right) = c\ . \tag{21}$$

We introduce a differentiable function $\hat{\ell}(a)$ such that $\forall(\zeta,a) \in \mathcal{S}$ and $t \in [0,T]$ we have:

$$|\zeta[p(\zeta + \ell(a), a) - p(\zeta + \hat{\ell}(a), a)]| \leq 1/4\alpha(c), \tag{22}$$

$$|\hat{\ell}(a) - \ell(a)| \leq d/2, \tag{23}$$

$$|\tilde{\beta}(|\zeta - \hat{\ell}(a)|, t) - \tilde{\beta}(|\zeta - \ell(a)|, t)| \leq \frac{d}{2} \tag{24}$$

and note that such a function always exist as we can approximate the continuous function $\ell(\cdot)$ with a differentiable function $\hat{\ell}(\cdot)$ to arbitrary accuracy on compact sets. Let $\epsilon_1^* > 0$ be such that

$$\epsilon_1^*|\zeta \cdot D\hat{\ell}(a)| \leq \frac{1}{4} \cdot \alpha(c)\ . \tag{25}$$

Moreover, let $\epsilon_2^* > 0$ be such that:

$$\beta_a(a_0, \epsilon_2^* T) > a^*\ , \tag{26}$$

where β_a comes from (11) and note that such a number always exists because $\beta_a \in \mathcal{KL}$. Finally, we let $\epsilon^* := \min\{\epsilon_1^*, \epsilon_2^*\}$.

Next, we show that the above constructed ϵ^* and $\tilde{\beta}$ satisfy conditions in our claim. In the rest of the proof we let $\epsilon \in (0, \epsilon^*)$ be arbitrary. Moreover, we introduce the change of time $\tau := \delta \cdot t$ and a transformation of coordinates $\zeta := x - \hat{\ell}(a)$ and rewrite (6) as follows:

$$\begin{aligned} \frac{d\zeta}{d\tau} &= a \cdot p(\zeta + \hat{\ell}(a), a) - \epsilon \cdot a \cdot D\hat{\ell}(a) \quad &(27) \\ \frac{da}{d\tau} &= -\epsilon \cdot g(a), \quad a(0) = a_0\ , &(28) \end{aligned}$$

where we have also used (9). We introduce the following Lyapunov function $V(\zeta) := \frac{1}{2}\zeta^2$ and taking its derivative along (27) we have from Propositions 2 and 3 and inequalities (22), (25) :

$$\begin{aligned} \frac{dV}{d\tau} &= \zeta[ap(\eta + \hat{\ell}(a), a) - \epsilon \cdot a \cdot D\hat{\ell}(a)] \\ &= a[\zeta p(\zeta + \ell(a), a) - \epsilon \cdot \zeta \cdot D\hat{\ell}(a)] \\ &\quad + a[\zeta(p(\zeta + \hat{\ell}(a), a) - p(\zeta + \ell(a), a))] \\ &\leq a[-\alpha(|\zeta|) + \frac{1}{2}\alpha(c)],\ \forall(\zeta, a) \in \mathcal{S}\ . \end{aligned} \tag{29}$$

The proof is completed by stating and proving several facts:

Fact 1: The set $\mathcal{S}_2$ is forward invariant.

Proof of Fact 1: This is straightforward from (29) and the fact that $a(\cdot)$ is monotonically decreasing.

Fact 2: For the number T defined by (21) we have that for any solution initialized at $|\zeta(0)| \leq D + d/2$ and $a(0) = a_0$ we have $(\zeta(T), a(T)) \in \mathcal{S}_2$.

Proof of Fact 2: If the initial state is in $\mathcal{S}_1$ we have nothing to prove since Fact 1 holds. Note that (29) guarantees that as long as we are in the set $\mathcal{S} - \mathcal{S}_1$, $\zeta(\cdot)$ is monotonically decreasing and we have:

$$|\zeta(\tau)| \leq \tilde{\beta}(|\zeta(0)|, \tau),\ \forall \tau \geq 0\ .$$

Note that $a(\cdot)$ is also monotonically decreasing. Moreover, we can use the contradiction and our choice of ϵ_2^* in (26) to show that all solutions initialized at $|\zeta(0)| \leq D + d/2$ and $a(0) = a_0$ stay in the set $\mathcal{S}_2$ for all $\tau \in [0,T]$. Finally, if we assume that for some $\zeta(0) \in \mathcal{S} - \mathcal{S}_1$ we have $|\zeta(\tau)| > c$ for all $\tau \in [0,T]$, this implies:

$$\tilde{\beta}(D + d/2, \tau) \geq |\zeta(\tau)| > c,\ \forall \tau \in [0,T]\ ,$$

which contradicts the choice of T in (21) and, hence, we have $(\zeta(T), a(T)) \in \mathcal{S}_2$.

Fact 3: If $|\zeta(0)| \leq D + d/2$ and $a(0) = a_0$, then solutions of the system (27), (28) satisfy

$$(\zeta(\tau), a(\tau)) \in \mathcal{S},\ \forall \tau \geq 0\ .$$

Proof of Fact 3: It follows trivially from Facts 1 and 2.

We now complete the proof of the theorem. From Fact 3, it follows that for all solutions initialized at $|\zeta(0)| \leq D + d/2$ and $a(0) = a_0$, the inequality (29) holds for all $\tau \geq 0$ along solutions. Moreover, Fact 2 says that for any such solution there exists $\tau_1 \in [0,T]$ such that $|\zeta(\tau)| > c$ for all $\tau \in [0, \tau_1)$ and $|\zeta(\tau_1)| \leq c$. Since $a(\tau) > a^*$ for all $\tau \in [0,T]$, from (29) we have that for all $\tau \in [0, \tau_1)$ the solutions of the system satisfy:

$$\frac{dV}{d\tau} \leq -\frac{a^*}{2} \cdot \alpha\left(\sqrt{2V}\right)\ ,$$

which implies:

$$|\zeta(\tau)| \leq \tilde{\beta}(|\zeta(0)|, \tau),\ \tau \in [0, \tau_1)\ . \tag{30}$$

On the other hand, for all $\tau \geq \tau_1$ we have from Fact 1 that $(\zeta(\tau), a(\tau)) \in \mathcal{S}_1$, which implies

$$|\zeta(\tau)| \leq c \leq d/2,\ \forall \tau \geq \tau_1\ . \tag{31}$$

Note now that from our choice of $\hat{\ell}$, we have that $|x(0) - \ell(a(0))| \leq D$ implies $|\zeta(0)| = |x(0) - \hat{\ell}(0)| \leq |x(0) -$

$\ell(a(0))| + |\hat{\ell}(a(0)) - \ell(a(0))| \leq D + \frac{d}{2}$ and this in turn implies (30) and (31). By adding and subtracting some terms to (30) and (31), we conclude that for all $\tau \in [0, \tau_1] \subseteq [0, T]$:

$$\begin{aligned}&|x(\tau) - \ell(a(\tau))| \leq \tilde{\beta}(|x(0) - \ell(a(0))|, \tau) + \\ &|\tilde{\beta}(|x(0) - \hat{\ell}(a(0))|, \tau) - \tilde{\beta}(|x(0) - \hat{\ell}(a(0))|, \tau)| + \\ &|\hat{\ell}(a(\tau)) - \ell(a(\tau))| \\ &\qquad\qquad \leq \tilde{\beta}(|x(0) - \ell(a(0))|, \tau) + d\end{aligned}$$

and for $\tau \geq \tau_1$ we have:

$$|x(\tau) - \ell(a(\tau))| \leq d/2 + |\hat{\ell}(a(\tau)) - \ell(a(\tau))| \leq d \quad (32)$$

Combining these last two bounds completes the proof by noting that $\tau = \delta t$.

B. Proof of Proposition 1

From (1), we have the following averaged system

$$\dot{x}_{av} = \frac{\delta a}{2}\left[Dh(x_{av}) + \frac{a^2}{8}D^3h(x_{av})\right] = \frac{\delta a}{2}p(x_{av}, a), \quad (33)$$

which is a 3^{rd}-order polynomial. Since (12) satisfies Assumption 1 and there are two maxima, it is not hard to see that there exist three different real roots of $Dh(x) = 0$, that we denote as r_1, r_2, r_3. Without loss of generality, we assume that $x^* = r_3$. We rewrite $p(x, a)$ in (33) as

$$p(x, a) = -c_1(x - r_1)(x - r_2)(x - r_3) + \frac{c_2 a^2}{8}(x - \frac{r_1+r_2+r_3}{3})$$

where c_1 and c_2 are positive constants. By a linear transformation $w = x + \frac{r_1+r_2+r_3}{3}$, we have

$$p(w, a) = -4\left(w^3 + \lambda(a) \cdot w + \lambda_1\right)$$

where $\lambda(a) \overset{\triangle}{=} -\frac{3}{16}\alpha_1^2 - \frac{1}{2}\alpha_2 + \frac{3}{4}a^2$ is a continuous function with respect to a and $\lambda_1 \overset{\triangle}{=} -\frac{1}{32}\alpha_1^3 - \frac{1}{4}\alpha_3 - \frac{1}{8}\alpha_2 \cdot \alpha_1$ is a constant. For all $i = 1, 2, 3$, the zeros of $p(z, a)$ can be represented as

$$r_i(a) \overset{\triangle}{=} \sqrt[3]{-\frac{\lambda_1}{2} + \sqrt{\Lambda(a)}p_i} + \sqrt[3]{-\frac{\lambda_1}{2} - \sqrt{\Lambda(a)}q_i} \quad (34)$$

where $\Lambda(a) \overset{\triangle}{=} \left(\frac{\lambda_1}{2}\right)^2 + \left(\frac{\lambda(a)}{3}\right)^3$, $p_1 = q_1 = 0$, $p_2 = q_3 = \omega$, $p_3 = q_2 = \omega^2$, and $\omega \overset{\triangle}{=} \frac{-1+\sqrt{3}i}{2}$. The following facts are apparent for those three zeros:
1) $r_i, i = 1, 2, 3$ and $\Lambda(a)$ are all continuous functions with respect to a; 2) $D\Lambda(a) = \frac{\lambda(a)^2}{3} \cdot \frac{3}{2}a \geq 0, \quad \forall a \geq 0$ indicates that $\Lambda(a)$ increases monotonically with respect to a for all $a \geq 0$; 3) When $\Lambda(a) < 0$, three distinct zeros are real. In other words, $\Lambda(0) < 0$. 4) There also exists a^* such that $\Lambda(a^*) = 0$ and three zeros are real, two of them are equal. 5) When $a \geq a^*$, $\Lambda(a) > 0$, there exists only one real zero. We show that item 2 of Assumption holds.

From the above discussion, it is obvious that $r_1(a)$ is always real for any $a > 0$. We can conclude that, there exists a continuous function $\ell(a) = r_1(a) : R_{\geq 0} \to R$, which is the zero of $p(x, a)$.

We show that $\ell(a) \in (\frac{r_1+r_2+r_3}{3}, r_3]$. The following proposition is needed to show the result.

Proposition 4: Assume that $r_1 < r_2 < r_3$ are the real zeros of $Dh(x)$ and $h(r_1) < h(r_3)$, where $h(\cdot)$ is defined (12). Then $r_2 < \frac{r_1+r_2+r_3}{3}$.
Proof: We denote a continuous function $g(z) \overset{\triangle}{=} \int_{r_1}^{r_3} -(\zeta - r_1)(\zeta - z)(\zeta - r_3)d\zeta$, which has the property $Dg(z) < 0$, $\forall z \in (r_1, r_3)$. Therefore $g(z)$ decreases monotonically for any $z \in (r_1, r_3)$. Note that $g(\frac{r_1+r_3}{2}) = 0$, $h(r_1) < h(r_3) \Rightarrow g(r_2) < 0$, the monotonicity of $g(z)$ implies that $r_2 < \frac{r_1+r_3}{2}$, i.e., $\frac{r_1+r_2+r_3}{3} - r_2 > 0$, which completes the proof. ■

The direct result of Proposition 4 is $h_3(\frac{r_1+r_2+r_3}{3}) < 0$, consequently, we have the following proposition.

Proposition 5: Assume that $r_1 < r_2 < r_3$ are the real zeros of $Dh(x)$ and $h(r_1) < h(r_3)$, where $h(\cdot)$ is defined (12). Then $\ell(a) \in (\frac{r_1+r_2+r_3}{3}, r_3]$.
Proof: Denote $A = (r_2, \frac{r_1+r_2+r_3}{3}]$, $B = (r_3, \infty)$ and $L = (\frac{r_1+r_2+r_3}{3}, r_3]$, for all $a \geq 0$, we have,

$$\begin{gathered}p(x, a) > 0 \quad \forall x \in A, \qquad p(x, a) < 0, \quad \forall x \in B, \\ p(\frac{r_1+r_2+r_3}{3}, a)p(r_3, a) < 0.\end{gathered}$$

Therefore, we know that, for any positive a, there is no real roots of $p(x, a)$ in any interval A or B while there always exists at least one real zero of $p(x, a)$ in the interval L.

When $a \geq a^*$, the uniqueness of $\ell(a)$ means that $\ell(a) \in (\frac{r_1+r_2+r_3}{3}, r_3], \forall a \geq a^*$. When $a < a^*$, the continuity of $\ell(a)$ implies that $\ell(a)$ can be in either A, or B or L. This immediately implies the conclusion of Proposition 5. ■

A direct consequence of Proposition 5 is $l(0) = r_3 = x^*$, since, r_3 is only real zero of $p(x, 0)$ in the interval L. This leads to the following proposition.

Proposition 6: Assume that $r_1 < r_2 < r_3$ are the real zeros of $Dh(x)$ and $h(r_1) < h(r_3)$, where $h(\cdot)$ is defined (12). Then $D_1 p(\ell(a), a) < 0$.

Proof: According to Proposition 5, $\ell(a)$ is in the interval L. Assume that $\exists x_a \in L$ such that $D_1 p(x_a, a) = 0$. Note $D_1^2 p(x_a, a) = \frac{a^2}{8}D^3h(x_a) < 0$, x_a is a local maximum of $p(x, a)$, i.e. $p(x_a + \lambda, a) < p(x_a, a)$ and $p(x_a - \lambda, a) < p(x_a, a)$, $\forall \lambda > 0$. However, it contradicts the fact that x_a is the real root of $p(x, a)$, which satisfies,

$$[p(x_a + \lambda, a) - p(x_a, a)]\,[p(x_a - \lambda, a) - p(x_a, a)] < 0,$$

Since $D_1 p(\ell(a), a) \neq 0, \forall \ell(a) \in L$ and $D_1 p(l(0), a) < 0$, we conclude that $D_1 p(\ell(a), a) < 0$ for all $a \geq 0$ ■