On the Choice of Dither in Extremum Seeking Systems: a Case Study

D. Nešić, Y. Tan, and I.M.Y.Mareels

The Department of Electrical & Electronics,
The University of Melbourne, Parkview, VIC 3010

Abstract

We discuss how the choice of dither (excitation signal) affects the performance of extremum seeking using a benchmark situation: a static scalar map; and a simple scalar extremum seeking scheme. Our comparisons are based on the performance of the system with different dithers in terms of three performance indicators: the speed of convergence, domain of attraction and accuracy (i.e. the ultimate bound on trajectories). Our analysis explicitly shows how the dither shape affects each of these performance indicators. Our study suggests that the practitioners using extremum seeking control should consider the dither shape as an important design parameter. Computer simulations support our theoretical findings.

1 Introduction

Extremum seeking (ES) control is a paradigm whose goal is to find an extremum value of an unknown nonlinear mapping. Although this method dates back to the early 1950's and 1960's, the first rigorous local stability analysis for an ES scheme was recently proved in (Krstić and Wang, 2000) and later extended to semi-global stability analysis in (Tan *et al.*, 2005; Tan *et al.*, 2006*a*). This has spurred a renewed interest in this research area (Ariyur and Krstić, 2003; Popović *et al.*, 2003; Guay and Zhang, 2003; Guay *et al.*, 2004; Peterson and Stefanopoulou, 2004) that has lead to numerous practical implementations of the scheme.

In our case study we focus on the following nonlinear static map

$$y = h(x), \tag{1}$$

where $h(\cdot)$ is not known, but it is known that $h(\cdot)$ has, for simplicity, a maximum $h(x^*)$. An ES mechanism will drive the output $y(t)$ to a small neighborhood of $h(x^*)$. In this paper, for simplicity, we focus on the scalar case, i.e., both y and x are scalar, that is the simplest version of the extremum seeking considered in (Tan *et al.*, 2006*a*).

In comparing the performance of the system with various dithers (excitation signals), we concentrate on three performance indicators: speed of convergence, domain of convergence and accuracy (i.e. the ultimate bound on trajectories). Some of our analysis is done using an

appropriate average system that approximates well the real closed loop behavior if certain parameters are chosen sufficiently small.

First, it is shown that an auxiliary gradient system plays a crucial role in quantifying the performance of the extremum seeker with different dithers. Indeed, it is shown that with arbitrary dither shape it is possible to tune the parameters in the controller and the dither amplitude to recover (with arbitrarily small error) the domain of attraction and the accuracy of the auxiliary gradient system. On the other hand, the closed loop system convergence speed depends on the convergence speed of the gradient system, as well as a scaling factor that is a product of four numbers: the amplitude of the dither, the frequency of the dither, a controller parameter and the power of the normalized dither (dither of the same shape with unit amplitude and period 2π). This bound is tight and it explicitly states how the dither shape affects the convergence speed. For instance, we can show that a square wave provides twice as fast a convergence as a sine wave of the same amplitude and frequency, if appropriate parameters are sufficiently small. The simulation results also show that the square-wave dither works better than sinusoidal dither which is used typically in the literature (Ariyur and Krstić, 2003). Hence, our results demonstrate that the dither shape is an important degree of freedom in the design and tuning of extremum seeking controllers and its choice should be given careful consideration.

Our results benefit from the scalar extremum seeking scheme that was introduced for the first time in (Tan *et al.*, 2006*a*). However, our main result improves upon results in (Tan *et al.*, 2006*a*) that were presented only for sinusoidal dither and that presented more conservative estimates than the ones derived in this paper. Indeed, we use a different proof technique from the one used in (Tan *et al.*, 2006*a*) in order to provide tighter stability estimates and to explicitly show how performance of the extremum seeking controller is affected by the choice of dither.

This paper is organized as follows. In Section 2 we present the benchmark plant and controller that we consider, as well as the standing assumptions. Main results is given in Section 3 followed by discussions and examples in Section 4. Conclusions are presented in Section 5.

2 A Benchmark Example

We denote the set of real numbers as R. Given a sufficiently smooth function $h : R \to R$, we denote its ith derivative as $D^i h(x)$. When $i = 1$ we write simply $Dh(x) := D^1(x)$.

We consider a static mapping $h(\cdot)$ with the first order extremum seeking controller (see Figure 1). This control law is a simplified version of the schemes considered in (Krstić and Wang, 2000) that was introduced in (Tan *et al.*, 2006*a*). In particular, this is the simplest possible variant of schemes considered in (Krstić and Wang, 2000) that can still illustrate that the choice of dither affects the extremum seeking in several different aspects. The model of the closed loop system in Figure 1 is given by:

$$\dot{x} = \delta \cdot \omega \cdot h(x + d(t)) \cdot d(t), \tag{2}$$

where $h : R \to R$ is sufficiently smooth. The signal $d(\cdot)$ is referred to as "dither" and $\delta > 0$ and $\omega > 0$ are parameters that the designer can choose.
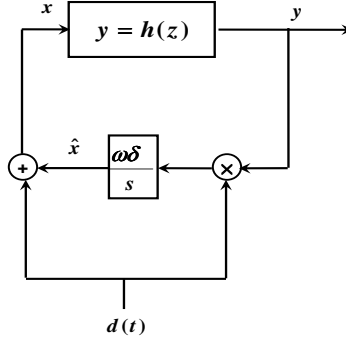
Figure 1: A peak seeking feedback scheme

We use the following assumptions:

Assumption 1 : *There exists a maximum x^* of $h(\cdot)$ such that*

$$Dh(x^*) = 0; \quad D^2h(x^*) < 0 . \tag{3}$$

Assumption 2 *Dither signals $d(\cdot)$ are periodic functions of period $T > 0$ (and frequency $\omega = \frac{2\pi}{T}$) that satisfy:*

$$\int_0^T d(s)ds = 0; \quad \frac{1}{T}\int_0^T d^2(s)ds > 0; \quad \max_{s \in [0,T]} |d(s)| = a;$$

where $a > 0$ is the amplitude of the dither.

The parameter ω in (2) is chosen to be the same as the frequency of the dither signal.

For comparison purposes, three special kinds of dither are used repeatedly in our examples: sine wave, square wave and triangle wave. The sine wave is defined in the usual manner. The square wave and triangle wave of unit amplitude and period 2π are defined as follows:

$$sq(t) \quad := \quad \begin{cases} 1, & t \in [2\pi k, \pi(2k+1)) \\ -1, & t \in [\pi(2k+1), 2\pi(k+1)) \end{cases}$$

$$tri(t) \quad := \quad \begin{cases} \frac{2}{\pi}(t - 2\pi k), & t \in [2\pi k, \pi(2k+\frac{1}{2})) \\ \frac{2}{\pi}(-t - \pi(2k-1)), & t \in [\pi(2k+\frac{1}{2}), \pi(2k+\frac{3}{2})) \\ \frac{2}{\pi}(t - \pi(2k+2)), & t \in [\pi(2k+\frac{3}{2}), 2\pi(k+1)) \end{cases}$$

Note that by definition, the signals $\sin(t)$, $sq(t)$ and $tri(t)$ are of unit amplitude and period 2π. We can generate similar signals of arbitrary amplitude a and frequency ω, e.g., $a \cdot sq(\omega t)$. We will often use the "power" (average of the square of the signal) of the normalized dithers (unit amplitude and period 2π):

$$P_{sq} = \frac{1}{2\pi}\int_0^{2\pi} sq^2(s)ds = 1; \quad P_{sin} = \frac{1}{2\pi}\int_0^{2\pi} \sin^2(s)ds = \frac{1}{2}; \quad P_{tri} = \frac{1}{2\pi}\int_0^{2\pi} tri^2(s)ds = \frac{1}{3} \tag{4}$$

3

In general, we use P_d to denote the power of dithers $d(\cdot)$ with amplitude equal to 1 and period 2π. The power for dither $d(\cdot)$ with amplitude $a \neq 1$ is equal to $a^2 P_d$. We emphasize that our results apply to arbitrary dithers satisfying Assumption 2.

Remark 1 *Assumption 2 is needed in our analysis that is based on averaging of (2). We note that most extremum seeking literature (see (Ariyur and Krstić, 2003)) uses dither signals of the form $d(t) = a\sin(\omega t)$ which obviously satisfy our Assumption 2.*

Introducing the coordinate change $\tilde{x} = x - x^*$, we can rewrite (2) as follows:

$$\dot{\tilde{x}} = \delta \cdot \omega \cdot h(\tilde{x} + x^* + d(t)) \cdot d(t) =: \delta\omega f(t, \tilde{x}, d). \tag{5}$$

It was shown in (Tan *et al.*, 2006*a*, Theorem 1) that under a stronger version of Assumption 1 (uniqueness of the maximum) and with the sinusoidal dither $d(t) = a\cdot\sin(\omega t)$, where $\omega = 1$ (that satisfies Assumption 2) we have that for any compact set $\mathcal{D}$ and any $\nu > 0$ we can chose the amplitude of the dither $a > 0$ and $\delta > 0$ and find a class $\mathcal{KL}$ function β, which depends on δ and $d(\cdot)$, such that the solutions of the closed loop system (5) satisfy:

$$|\tilde{x}(t_0)| \in \mathcal{D} \implies |\tilde{x}(t)| \leq \beta\left(|\tilde{x}(t_0)|, t - t_0\right) + \nu, \tag{6}$$

for all $t \geq t_0 \geq 0$.

Note that $\mathcal{D}$, ν and β are *performance indicators* since they quantify different aspects of the performance of the extremum seeking algorithm. We will show later that each of these indicators is affected by our choice of dither $d(\cdot)$ and the parameter $\delta > 0$. In particular, we have that:

- **Speed of convergence** of the algorithm is captured by the function β. Obviously, we would like convergence to be as fast as possible.

- **Domain of convergence** is quantified by the set $\mathcal{D}$. In particular, we would like to make the domain of convergence (attraction) as large as possible.

- **Accuracy** of the algorithm is quantified by the number $\nu > 0$ since all trajectories starting in the set $\mathcal{D}$ eventually end up in the ball B_ν, where we have that $|x(t) - x^*| \leq \nu$. Indeed, the smaller the number ν, the closer we eventually converge to the maximum x^* (hence, the accuracy of the algorithm is better).

It turns out that a direct analysis of the system (5) to estimate $\mathcal{D}, \nu, \beta$ is hard but the system can be analyzed via an appropriate auxiliary averaged system. We will carry out such an analysis in the next section.

3 Main results

In this section, we present the main result (Theorem 1) that describes in detail how different dithers affect the domain of attraction and speed of convergence, as well as the accuracy of extremum seeking control. It is shown that the square wave produces the fastest convergence among all signals with the same amplitude and frequency, if the amplitude a and the

parameter δ in the controller are sufficiently small. Moreover, it is shown that in the limit as the amplitude is reduced to zero, all dithers yield almost the same domain of attraction and accuracy.

Consider the following auxiliary gradient system:

$$\dot{\zeta} = Dh(\zeta + x^*) \ . \tag{7}$$

Because of Assumption 1, the system (7) has the property that x^* is an asymptotically stable equilibrium[1]. Let $\mathcal{D}$ denote the domain of attraction of x^* for the system (7) and note that since $h(\cdot)$ is assumed smooth, the set $\mathcal{D}$ is a neighborhood of x^*. In other words, a consequence of Assumption 1 is that there exists $\beta \in \mathcal{KL}$ and a set $\mathcal{D}$ such that for all $t \geq 0$ the solutions of (7) satisfy:

$$\zeta_0 \in \mathcal{D} \quad \Rightarrow \quad |\zeta(t)| \leq \beta(|\zeta_0|, t) \tag{8}$$

Using this auxiliary system, we can state our main result:

Theorem 1 *Suppose that Assumption 1 holds and consider the closed loop system (2) with an arbitrary dither $d(\cdot)$ for which Assumption 2 holds, where $a > 0$ is the dither amplitude. Let $\mathcal{D}$ and β come from (8). Then, for any strict compact subset $\hat{\mathcal{D}}$ of $\mathcal{D}$ and any $\nu > 0$, there exists $a^* > 0$ and $\delta^* > 0$ such that for any $a \in (0, a^*]$, $\delta \in (0, \delta^*]$ and any $\omega > 0$ we have that solutions of (2) satisfy:*

$$\tilde{x}_0 \in \hat{\mathcal{D}} \Rightarrow |\tilde{x}(t)| \leq \beta\left(|\tilde{x}_0|, \delta\omega a^2 P_d(t - t_0)\right) + \nu \tag{9}$$

Sketch of proof of Theorem 1: Consider the system (2). First, let $\omega > 0$ be arbitrary frequency of the dither and introduce the change of time scale $\tau := \omega t$. Then, we can rewrite (2) as follows:

$$\frac{d\tilde{x}}{d\tau} = \delta \cdot h\left(\tilde{x} + x^* + d\left(\frac{\tau}{\omega}\right)\right) \cdot d\left(\frac{\tau}{\omega}\right) \ ,$$

where $d(\frac{\cdot}{\omega})$ has period 2π in time scale τ. We apply the Taylor series expansion for $h\left(\tilde{x} + x^* + d\left(\frac{\tau}{\omega}\right)\right)$ around $\tilde{x} + x^*$ and using averaging method in (Sanders and Verhulst, 1985; Tan *et al.*, 2006*a*), we obtain the average system in the form:

$$\frac{d\tilde{x}}{d\tau} = \delta \cdot a^2 \cdot P_d \cdot Dh(\tilde{x} + x^*). \tag{10}$$

Using a transformation of coordinates given in (Tan *et al.*, 2006*a*), we can show that the actual system (2) can be regarded as the average system (10) that is additively perturbed with the regular perturbations that can be reduced arbitrarily by reducing δ and a simultaneously. Hence, we can show appropriate closeness of solutions on compact time intervals between the actual and average systems. We can use the trajectory based proofs for stability via averaging (Teel *et al.*, 1999) to complete the proof.

---

[1]Moreover, all local maxima of $h(\cdot)$ are asymptotically stable equilibria of (7) and all local minima of $h(\cdot)$ are unstable equilibria of (7).

Remark 2 *We emphasize that the auxiliary gradient system (7) plays a crucial role in terms of achievable performance of the extremum seeking controller. Indeed, $\mathcal{D}$ and β are* independent *of the choice of dither and Theorem 1 specifies how they affect the achievable domain of attraction $\hat{\mathcal{D}}$ of the closed loop (2), as well as the speed of convergence via the function β.*

Remark 3 *We now discuss Theorem 1 in more detail to explain how dither shape affects the domain of attraction, accuracy and convergence speed of the closed loop system. We note that the controller parameter (a, δ) needs to be tuned appropriately in order for Theorem 1 to hold.*

Domain of attraction: It is shown that any dither satisfying Assumption 2 can yield a domain of attraction $\hat{\mathcal{D}}$ that is an arbitrary strict subset of the domain of attraction of the gradient system (7) if a and δ are sufficiently small. We emphasize that $\omega > 0$ can be arbitrary and a and δ do not depend on it.

Accuracy: The ultimate bound that is quantified by the number ν can be made arbitrarily small by any dither satisfying Assumption 2 if a and δ are sufficiently small. Hence, in the limit, all dithers perform equally well in terms of domain of attraction and accuracy.

Convergence speed: We emphasize that β in (9) is the same as β in (8) for any dither $d(\cdot)$. The main difference in speed of convergence comes from the scaling factor within the function β:

$$\delta \cdot \omega \cdot a^2 \cdot P_d \, , \tag{11}$$

where δ is a controller parameter, a and ω are respectively the amplitude and frequency of dither and P_d is the power of the normalized dither (with unit amplitude and period 2π). Note also that $\omega\delta$ is the integrator constant in Figure 1. Also, note that Theorem 1 holds for sufficiently small a and δ that are independent of ω which is an arbitrary positive number. Obviously, if the product (11) is larger than 1 then the closed loop system (2) converges faster than the auxiliary gradient system (7). Similarly, if the product (11) is smaller than 1, the system (2) converges slower than the gradient system (7).

The first observation is that for sufficiently small a and δ the bound in Theorem 1 holds for any ω. Hence, for fixed a, δ and P_d we have that the larger the ω, the faster the convergence. In other words, Theorem 1 shows that in our case study we can achieve arbitrarily fast convergence of the extremum seeking closed loop by making ω sufficiently large. Simulations in Example 1 verify our analysis. We also emphasize that this result is in general not possible to prove for general dynamic plants. For instance, the results in (Tan et al., 2006a) that are stated for general dynamical systems provide a similar bound as in (9) under the stronger assumption that ω is sufficiently small.

Suppose now that a, δ and ω are fixed and we are only interested in how the shape of dither affects the convergence. As we change dither, its (normalized) power P_d changes and as we can see from (4) that the square wave will yield twice larger normalized power than the sine wave and three times larger power than the triangle wave. Consequently, we can expect twice faster convergence with the square wave than with the sine wave and three times faster convergence than with the triangle wave. Simulation results in Example 2 that we present in the sequel are consistent with the above analysis.

Remark 4 *A weaker version of Theorem 1 was proved in (Tan et al., 2006a) for the sine wave dither only. Indeed, the results in (Tan et al., 2006a) do not consider arbitrary dither and the domain of attraction and convergence estimates are not as sharp as in Theorem 1. For instance, the relationship of convergence rate and the domain of attraction to the auxiliary system (7) was not shown in (Tan et al., 2006a) as this was impossible to do using the Lyapunov based proofs used in this reference. On the other hand, using the trajectory based proofs adopted in this paper, we can prove tight estimates as outlined in Theorem 1. Moreover, in (Tan et al., 2006a) it was not clear how the dither power P_d affects the convergence rate of the average system. Note that the function β in (9) is the same for any dither satisfying Assumption 2 and the only difference comes the parameters in (11) . However, the values of a^* and δ^* are typically different for different dithers.*

Remark 5 *We note that one can state and prove a more general version of Theorem 1 that applies to general dynamical plants and in this case $h(\cdot)$ is an appropriate reference-to-output map. With extra assumptions on the plant dynamics, one can use singular perturbation theory to prove this more general result (see for instance (Tan et al., 2006a) for a Lyapunov based proof in the case of sine wave dither). However, in this case we will need to require that ω is sufficiently small.*

The following proposition is obvious and it states that the power of the normalized square wave is larger than or equal to the power of any other normalized dither satisfying Assumption 2. In other words, for fixed δ and a for which (9) holds, the square wave is guaranteed to produce the fastest convergence over all dithers with the same amplitude and frequency.

Proposition 1 *Consider arbitrary $d(\cdot)$ satisfying Assumption 2. Then, we have that the power of the normalized dither satisfies:*

$$0 < P_d \leq P_{sq} = 1 .$$

4   Discussions and Examples

It has been shown in Theorem 1 that the convergence speed of the ES systems depends on the choice the dither shape P_d, amplitude a and frequency ω as well as δ. It also is shown that the domain of the attraction and accuracy of all dithers are almost the same as $\begin{cases} a \to 0 \\ \delta \to 0 \end{cases}$.

In this part, we use examples to illustrate various behaviors and simulations to confirm our theoretical findings. Furthermore, we also provide examples which illustrate that for a fixed non-zero amplitude different dithers do produce different domains of convergence and accuracy. Our results should motivate the users of extremum seeking control to experiment with different dithers in order to achieve the desired trade-off between convergence, domain of attraction or accuracy.

The first example illustrates that increasing the frequency of dither while keeping a, δ and P_d the same yields faster convergence.

Example 1 *Consider the quadratic mapping*

$$h(x) = -(x+4)^2 \qquad (12)$$

where $Dh(\tilde{x} + x^*) = -2\tilde{x}$. It is trivial to see that in this case $\mathcal{D} = R$ and $\beta(s,t) = se^{-2t}$ (see Theorem 1). The dither is chosen to be $d(t) = a\sin(\omega t)$. Hence, from (4) we have $P_{sin} = 1/2$. The initial condition is chosen as $x_0 = -2$. When we fix $a = 0.5$ and $\delta = 0.1$, the output response with different frequencies is shown in Figure 2. It is clear that the larger the ω is, the faster the convergence is.
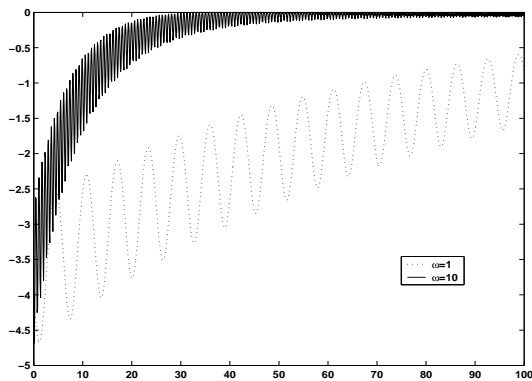

Figure 2: The output of the ES with different frequencies

Next, we will discuss the situation when $h(\cdot)$ is a quadratic mapping, where we assume $\omega = 1$. Consider the simplest possible case of quadratic maps:

$$h(x) = -x^2 + a_1 x + a_0 \ ,$$

we have that

$$Dh(x) = -2x + a_1$$

and since $x^* = \frac{a_1}{2}$, we can write with $\tilde{x} := x - x^*$:

$$Dh(\tilde{x} + x^*) = -2\tilde{x} \ .$$

Hence, the auxiliary gradient system (7) takes the following form:

$$\dot{\zeta} = -2\zeta \ .$$

It is not hard to show that in this case for arbitrary dither $d(\cdot)$ satisfying Assumption 2 we have that the average system is of the form:

$$\dot{\bar{x}} = -2a^2 P_d \delta \tilde{x} \ . \qquad (13)$$

Hence, in this special case we have that the average system for any dither satisfying Assumption 2 is globally exponentially stable. Indeed, for square wave, sine wave and triangular

8

wave we have from (4) that the following holds for all $x_0 \in R$, $t \geq 0$, respectively:

$$sq : \; |\tilde{x}(t)| \;=\; \exp\left(-2a^2\delta t\right)|\tilde{x}_0|$$
$$sin : \; |\tilde{x}(t)| \;=\; \exp\left(-a^2\delta t\right)|\tilde{x}_0|$$
$$tri : \; |\tilde{x}(t)| \;=\; \exp\left(-\frac{2}{3}a^2\delta t\right)|\tilde{x}_0| \; .$$

The square wave produces the fastest speed of convergence for the average system among all dithers with the same amplitude. The same can be concluded for the actual system using the proof of Theorem 1. This suggests that the square-wave dither should be the prime candidate to use in the ES system for fast convergence speed, although this dither is rarely considered in the literature (Ariyur and Krstić, 2003). Indeed, all references that we are aware of use a sinusoidal dither signal. The simulation results shown in Example 2 illustrates that the convergence speed of ES with the square wave is fastest among all dithers with the same amplitude.

Example 2 *The simulation is done for the following system where* $\omega = 1$:

$$\dot{x} = \delta h(x + d(t))d(t)$$

where $h(x) = -(x+4)^2$. *In the new coordinate,* $\tilde{x} = x - x^* = x + 4$, *we have*

$$\dot{\tilde{x}} = \delta h(\tilde{x} + x^* + d(t))d(t) = -\delta(\tilde{x} + d(t))^2 d(t).$$

The averaged system is given in (13). The simulation result is shown in Figure 3, where $a = 0.1$ *and* $\delta = 0.5$. *Simulations show that the extremum seeking controller with the square wave dither converges fastest.*
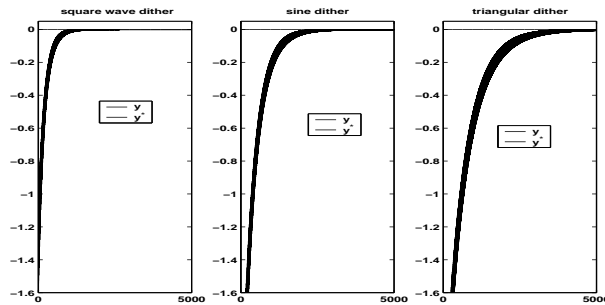

Figure 3: The performance of the extremum seeking schemes of different excitation signal

Finally, we investigate that the domain of the attraction and accuracy of the ES when a is not small enough. It should be noted that when using the averaging method in Theorem 1, the higher order terms in $d(t)$ do not appear in the averaged system (10). It indicates that the trajectory of the averaged system (10) is close to the trajectory of the actual system (2) only when both a and δ are sufficiently small. In order to investigate the performance

of the ES with different dithers with arbitrary amplitude, we use the standard averaging method (KHalil, 2002, Chpater 10) which averages the obtained time varying function over the interval $[0, T]$.

We note that for sufficiently smooth (or analytic) maps $h(\cdot)$ we can use the Taylor series expansion to obtain the following averaged system (by standard averaging) for a general dither that

$$
\begin{aligned}
f_{av}^d(\tilde{x}, a) &= \delta \cdot \frac{1}{T} \int_0^T f(\tau, \tilde{x}, d(\tau)) d\tau \\
&= \delta \cdot \left(\sum_{i=1}^N c_i^d \cdot D^{2i-1} h(\tilde{x} + x^*) + R_N^d(\tilde{x} + x^*, a) \right)
\end{aligned}
\tag{14}
$$

where $f(\cdot, \cdot, \cdot)$ is defined in (5), R_N^d is the reminder and c_i^d is defined as follows

$$
c_i^d := \frac{1}{(2i-1)!T} \int_0^T d^{2i}(s) ds .
\tag{15}
$$

These two averaging methods yield different approximation for the same function $f(\cdot, \cdot, \cdot)$. When both a and δ are sufficiently small, these two methods would produce the same results. However, the standard averaging method provides a better approximation when a is not small enough.

In particular, for polynomials maps $h(\cdot)$ we have that there exists sufficiently large N such that $D^i h = 0$ for all $i \geq N$ and hence $R_N = 0$. Similarly, for more general analytic maps $h(\cdot)$ we can write the average system as an infinite sum involving derivatives of h. Different dither yields different average system which is a weighted sum of the derivatives of $h(\cdot)$. In particular, for a square wave, sine wave and triangle wave with amplitude a and period 2π direct calculations yield:

$$
c_1^{sq} = a^2, \quad c_2^{sq} = \frac{a^4}{6}; \quad c_1^{sin} = \frac{a^2}{2}, \quad c_2^{sin} = \frac{a^4}{16}; \quad c_1^{tri} = \frac{a^2}{3}, \quad c_2^{tri} = \frac{a^4}{30}
\tag{16}
$$

Obviously the choice of the dither does affect the performance of the ES system. Next we discuss in more detail 4th order polynomials (quartic mappings), where we assume $\omega = 1$. For a 4th order polynomial, using a standard averaging method, we can write in general that:

$$
f_{av}^d(\tilde{x}, a) = c_1^d D h(\tilde{x} + x^*) + c_2^d D^3 h(\tilde{x} + x^*) ,
\tag{17}
$$

where $c_i^d, i = 1, 2$ is defined in (15). Moreover, for the square wave, sine wave and triangle wave these coefficients are given in (16).

Stability properties of the closed loop system (2) crucially depend on the properties of the average system. Moreover, the above formulas specify how the average system depends on different dither and they illustrate the flexibility that may be gained through the choice of the dither signal. We emphasize that for any dither, the average system can be written in the form (10) which indicates that for sufficiently small a and δ, the extremum seeking controller operates in almost the same way as classical "gradient descent" methods.

Example 3 *In this example, we show that the square wave may yield inferior accuracy compared to the sine wave and the triangle wave dithers with the same amplitude a, especially for large values of a. We use signals $d_1(t) = a\,sq(t)$, $d_2(t) = a\sin(t)$ and $d_3(t) = a\,tri(t)$ of period 2π. Consider the following 4th order polynomial:*

$$h(x) = -x^4 - \tfrac{16}{3}x^3 + 22x^2 + 120x + 1 \ ,$$

that has a global maximum at $x = 3$, a local maximum at $x = -5$ and a local minimum at $x = -2$. The plot of this function is given in Figure 4.
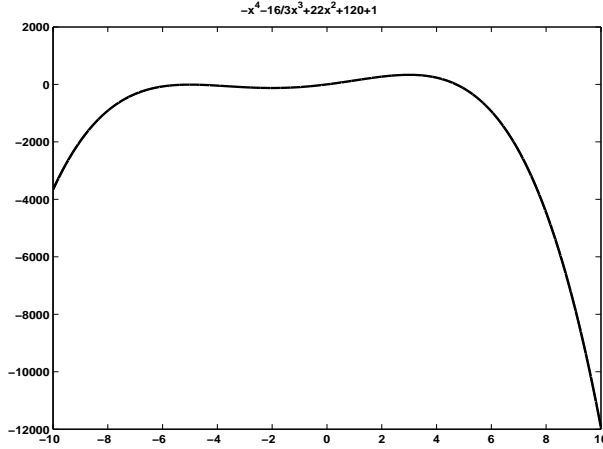

Figure 4: The plot of a 4th order polynomial.

Using (16) and (17), we obtain the following average systems with different dithers:

$$
\begin{aligned}
f_{av}^{d_1}(x, a) &= a^2(-4x^3 - 16x^2 + 44x + 120) + \tfrac{a^4}{6}(-24x - 32) \\
f_{av}^{d_2}(x, a) &= \tfrac{a^2}{2}(-4x^3 - 16x^2 + 44x + 120) + \tfrac{a^4}{16}(-24x - 32) \\
f_{av}^{d_3}(x, a) &= \tfrac{a^2}{3}(-4x^3 - 16x^2 + 44x + 120) + \tfrac{a^4}{30}(-24x - 32) \ .
\end{aligned}
$$

Since for each fixed $a > 0$ we are dealing with scalar systems of the form:

$$\dot{\tilde{x}} = f_{av}^{d_i}(\tilde{x}, a) \ ,$$

with $i = 1, 2, 3$, it is possible to precisely determine the domain of convergence and the accuracy of the average systems with different dithers. Indeed, consider the following equation

$$f_{av}^{d_i}(x, a) = 0 \ ,$$

for different dithers. This equation defines a bifurcation diagram that explains how the equilibria of the averaged system changes as a is varied. Moreover, the bifurcation diagram partitions the state space into regions on which $f_{av}^{d_i}(x, a)$ is either positive or negative. In Figure 5, we plotted the bifurcation diagrams for three different dithers with varying amplitude. For instance, the dotted lines in Figure 5 represent the bifurcation diagram of $f_{av}^{d_1}(x, a)$. Moreover, the "+" sign in Figure 5 indicates that in the area between the two curves we have

that $f_{av}^{d_1}(x, a) > 0$ and the two signs "−" indicate the areas where $f_{av}^{d_1}(x, a) < 0$. Hence, with the aid of the bifurcation diagram we can analyze the dynamics of the averaged system for each fixed a for any of the three dithers. It is obvious that for small a we have that the system has attractive equilibria located in the neighborhood of $x = 3$ and $x = -5$ (respectively the global and local maxima of $h(\cdot)$) and an unstable equilibrium in the neighborhood of $x = -2$ (the local maximum of $h(\cdot)$). However, the bifurcation diagram indicates that for $a = 4$ we have that all solutions of the average system with the square wave will converge to a neighborhood of $x = 1$ whereas all solutions of the average system with the triangle wave will converge to a neighborhood of $x = 1.8$. Hence, for this amplitude, the extremum seeking controller with a triangle dither provides a better accuracy than the same controller with the square wave dither. Simulation results also show that when $a = 4$, the better accuracy is obtained for the ES with a triangle dither.
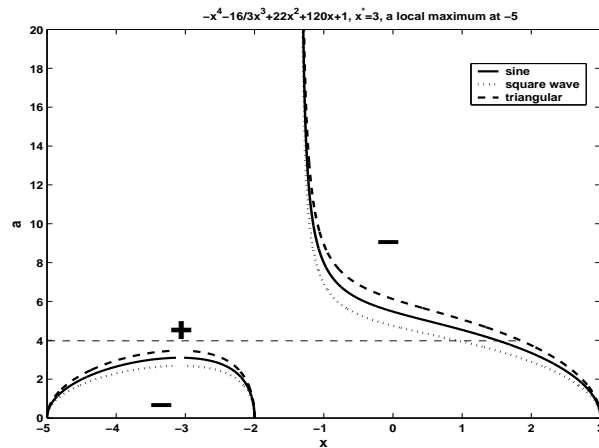

Figure 5: The bifurcation diagram for different dithers with the same amplitude.

Example 4 *We revisit the system in Example 3 but in this case we consider dithers that have the same powers, that is we use signals $d_1(t) = a sq(t)$, $d_2(t) = a\sqrt{2}\sin(t)$ and $d_3(t) = a\sqrt{3}tri(t)$ with period 2π. It is easy to see that the power of all three dithers is $p = a^2$ for any a. The reason for this comparison is that in this case we have that all three dithers yield the same convergence speed for sufficiently small a (since in (17) we have $c_1^{d_1} = c_1^{d_2} = c_1^{d_3} = a^2$) and we want to analyze the differences in accuracy and domain of attraction as we vary the power p of the signals (see Figure 6). In this case, average systems for different dithers are given by:*

$$
\begin{aligned}
f_{av}^{d_1}(x, p) &= p(-4x^3 - 16x^2 + 44x + 120) + \tfrac{p^2}{6}(-24x - 32) \\
f_{av}^{d_2}(x, p) &= p(-4x^3 - 16x^2 + 44x + 120) + \tfrac{p^2}{4}(-24x - 32) \\
f_{av}^{d_3}(x, p) &= p(-4x^3 - 16x^2 + 44x + 120) + \tfrac{3p^2}{10}(-24x - 32) \ .
\end{aligned}
$$

Using analysis similar to the one presented in Example 3, it can be seen from Figure 6 that the accuracy of extremum seeking to the global maximum $x = 3$ is better with the square wave dither than with the sine wave or the triangle wave that have the same power. This

12

is opposite from the observations made in Example 3 for signals with the same amplitude. Simulation results support our observations.
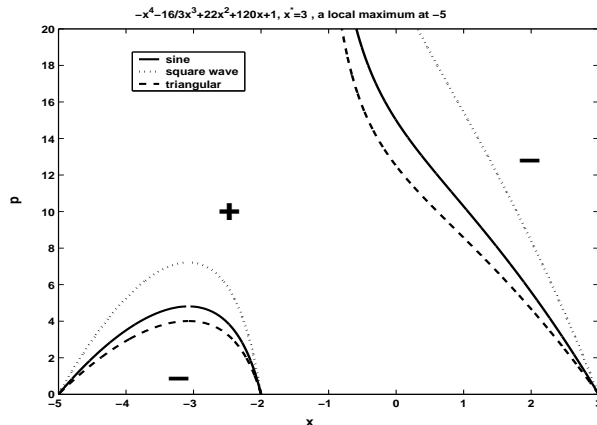

Figure 6: The bifurcation diagram for different dithers with the same power.

Remark 6 *If the mapping $h(\cdot)$ was known to the designer, it would be easy to verify which dither performs best in terms of accuracy or domain of attraction by plotting the bifurcation diagrams such as the ones given in Examples 3 and 4. Hence, it may seem that we have all the tools needed to make appropriate dither selection in general. However, the underlying assumption in extremum seeking control is that $h(\cdot)$ is unknown and, hence, the only way to verify which dither performs better is through experimenting.*

Remark 7 *The bifurcation diagrams presented in Figure 5 can be used to develop another extremum seeking scheme that achieves global extremum seeking in the presence of local extrema. It was shown in (Tan et al., 2006b) under very general conditions that this strategy yields global convergence under certain assumptions on the bifurcation diagram. In particular, it is required that there exists a solution $\ell(a)$ to the equation $f_{av}(x, a) = 0$, which is continuous, it is unique for large a and such that $\ell(0) = x^*$. Note that all three bifurcation diagrams in Figure 5 satisfy these conditions and, hence, global extremum seeking is possible with the scheme from (Tan et al., 2006b). The idea is to initially choose a sufficiently large amplitude of the excitation signal and then reduce it to zero sufficiently slowly. We note that results in (Tan et al., 2006b) were derived for the sine wave dither only, whereas it is clear from this paper that any other dither can be used instead.*

5   Conclusions

We have presented results that illustrate how the choice of dither affects the performance of an extremum seeking scheme. Our results demonstrate that the dither is an important design parameter to be considered when tuning the extremum seeking controller. For instance, it is shown that for small amplitudes the square wave provides the best convergence rate among all dithers of the same amplitude and frequency. Our examples illustrate further the flexibility that may be gained through dither design and simulations support our theoretical findings.

13

References

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

Guay, M. and T. Zhang (2003). Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties. *Automatica* **39**, 1283–1293.

Guay, M., D. Dochain and M. Perrier (2004). Adaptive extremum seeking control of continuous stirred tank bioreactors with unknown growth kinetics. *Automatica* **40**, 1645–1650.

KHalil, H. K. (2002). *Nonlinear Systems*. 3 ed.. Prentice Hall, Upper Saddle River, New Jersey.

Krstić, M. and H. H. Wang (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica* **36**, 595–601.

Peterson, K. S. and A. G. Stefanopoulou (2004). Extremum seeking control for soft landing of an electromechanical valve actuator. *Automatica* **40**, 1063–1069.

Popović, D., M. Janković, S. Manger and A.R.Teel (2003). Extremum seeking methods for optimization of variable cam timing engine operation. In: *Proc. American Control Conference*. Denver, Colorado, USA.

Sanders, J. A. and F. Verhulst (1985). *Averaging Methods in Nonlinear Dynamical Systems*. Springer-Verlag, New York Inc.

Tan, Y., D. Nešić and I.M.Y. Mareels (2005). On non-local stability properties of extremum seeking control. In: *Proc. 16th IFAC World Congress*. Prague, Czech Republic.

Tan, Y., D. Nešić and I.M.Y. Mareels (2006*a*). On non-local stability properties of extremum seeking control. *will appear inAutomatica*.

Tan, Y., D. Nešić, I.M.Y. Mareels and A. Astolfi (2006*b*). On global extremum seeking in the presence of local extrema. Submitted to Conf. Decis. Contr., San Diego, USA, 2006.

Teel, A. R., J. Peuteman and D. Aeyels (1999). Semi-global practical asymptotic stability &averaging. *Systems &Control Letters* **37**, 329–334.

Teel, A. R., L. Moreau and D. Nešić (2003). A unified framework for input-to-state stability in systems with two time scales. *IEEE Transactions on Automatic Control* **48**, 1526–1544.