Author/s:
Nesic, D;Thang, N;Tan, Y;Manzie, C

Title:
A non-gradient approach to global extremum seeking: An adaptation of the Shubert algorithm

Date:
2013-03-01

Citation:
Nesic, D., Thang, N., Tan, Y. & Manzie, C. (2013). A non-gradient approach to global extremum seeking: An adaptation of the Shubert algorithm. AUTOMATICA, 49 (3), pp.809-815. https://doi.org/10.1016/j.automatica.2012.12.004.

Persistent Link:
https://hdl.handle.net/11343/299555

A non-gradient approach to global extremum seeking: an adaptation of the Shubert algorithm ⋆

Dragan Nešić [a], Thang Nguyen [a], Ying Tan [a], Chris Manzie [b]

[a] *The Department of Electrical & Electronics Engineering, The University of Melbourne, Melbourne, VIC 3010, Australia*

[b] *The Department of Mechanical Engineering, The University of Melbourne, Melbourne, VIC 3010, Australia*

Abstract

The main purpose of this paper is to adapt the so-called Shubert algorithm for extremum seeking control of general dynamic plants. This algorithm is a good representative of the "sampling optimization methods" that achieve global extremum seeking on compact sets in presence of local extrema. The algorithm applies to Lipschitz mappings; the model of the system is assumed unknown but the knowledge of its Lipschitz constant is assumed. The controller depends on a design parameter, the "waiting time" and tuning guidelines that relate the design parameter and the region of convergence and accuracy of the algorithm are presented. The analysis shows that semi-global practical convergence (in the initial states) to the *global extremum* can be achieved in presence of local extrema if compact sets of inputs are considered. Numerical simulations for global optimization in presence of local extrema are provided to demonstrate the proposed approach.

Key words: Extremum seeking control; Global optimization; Shubert's algorithm; Lipschitz.

1 Introduction

Most extremum seeking controllers in the literature are based on optimization methods that require the derivatives of an unknown steady-state input-output map to be estimated online, see (Ariyur and Krstić, 2003; Tan *et al.*, 2006; Moase *et al.*, 2010; Teel and Popović, 2001); moreover, they typically find local extrema only. On the other hand, optimization methods based on "sampling techniques" may not require the derivatives of the map and they find a global extremum on a compact set even in the case when local extrema exist, see (Strongin and Sergeyev, 2000; Shubert, 1972; Hansen *et al.*, 1992*a*). We are not aware of any extremum seeking methods based on these sampling techniques. This paper opens opportunities for further research on adapting other sampling techniques in (Strongin and Sergeyev, 2000; Hansen *et al.*, 1992*a*) for extremum seeking control.

The main purpose of this paper is to adapt one such algorithm - the so-called Shubert algorithm - for extremum seeking control of dynamic systems. A periodic sampled-data controller is used in order to apply the discrete-time Shubert algorithm to the continuous-time plant. Similar to (Teel and Popović, 2001), the tuning parameter "waiting time" is introduced. It is shown that the closed loop system with the new extremum seeking algorithm can approximately achieve global optimization [1] of the steady-state input output map in presence of local extrema. More precisely, we show semi-global [2] practical convergence with respect to a controller design parameter, the so-called waiting time T. The presented results apply to a general class of dynamic plants. Our proofs can be interpreted as an appropriate robustness analysis of the Shubert algorithm given in (Shubert, 1972); in particular, a result proved in (Shubert, 1972) follows directly from our analysis.

The paper is organized as follows. Section 2 revisits the Shubert algorithm and in Section 3 we show how it can be adapted for extremum seeking control. The main result

⋆ Corresponding author Y. Tan. This research is supported by the Australian Research Council under the Discovery Grant DP0985388

Email addresses: dnesic@unimelb.edu.au (Dragan Nešić), thangnthn@gmail.com (Thang Nguyen), yingt@unimelb.edu.au (Ying Tan), manziec@unimelb.edu.au (Chris Manzie).

[1] Whenever we talk about global optimization of a given map, we assume that the set of its arguments is compact.

[2] Semi-global is with respect to the initial conditions of the dynamic plant. Precise definitions and statements of our results will be given later.

Preprint submitted to Automatica *17 October 2012*

Preprint submitted to Automatica
Received October 16, 2012 13:44:54 PST

is stated in Section 4. A numerical example is provided in Section 5. Conclusions are presented in Section 6.

2 Revision of the Shubert algorithm

First, we present the main assumptions and an algorithm from (Shubert, 1972) that can be used for global optimization of single-input-single-output (SISO) static systems. Note that this algorithm does not require the knowledge of the system model to perform the optimization and, hence, it is ideally suited for online optimization such as extremum seeking. Some knowledge of the plant model is assumed; indeed, the algorithm uses *the Lipschitz constant* of the model and the output measurements to construct a sequence of inputs that converge to the global extremum.

Note that (Shubert, 1972) considers the problem of finding the maximum of a map $Q(\cdot)$ on a compact interval $[a,b]$. The problem is viewed as an online optimization (extremum seeking) problem for the discrete-time static SISO system (1) because the main result in the next section generalizes this problem formulation to dynamic plants. Results of (Shubert, 1972) directly apply to static SISO plants (1) and they are recalled in this section. The main results in the next section demonstrate how the Shubert algorithm can be applied to dynamic plants in an extremum seeking fashion.

Consider the model of a discrete-time static SISO system [3]

$$y_k = Q(u_k)\ , \qquad k = 1, 2, \ldots, \tag{1}$$

where $y \in \mathbb{R}$ and $u \in [a,b]$ are respectively the output and input of the system. Consider the system (1) on a compact interval $[a,b]$ with $a < b$; that is, $Q : [a,b] \to \mathbb{R}$.

In order to apply the Shubert algorithm, the following assumption is needed.

Assumption 1 *The mapping $Q(\cdot)$ satisfies the Lipschitz condition with a positive constant L:*

$$|Q(u_1) - Q(u_2)| \le L|u_1 - u_2|, \quad \forall u_1, u_2 \in [a,b]. \tag{2}$$

Moreover, we assume that L is known and available to the designer whereas $Q(\cdot)$ is not.

Remark 1 *In practice, the Lipschitz constant of $Q(\cdot)$ may need to be estimated, see for example, (Strongin, 1973; Meewella and Marne, 1988; Hansen* et al.*, 1992c). Denote the minimum value of L that satisfies (2) as $L_{\min}$. In order to ensure that the Shubert algorithm works, it is necessary that the estimated Lipschitz constant $\widehat{L}$ satisfies $\widehat{L} \ge L_{\min}$. On the other hand, the larger the estimate $\widehat{L}$, the slower convergence of the algorithm (Shubert, 1972).*

Remark 2 *For simplicity of presentation, this note focuses on SISO systems to which SISO Shubert algorithm can be directly applied. There exist algorithms that extend the Shubert algorithm to multiple inputs, see for example, (Mladineo, 1986; Meewella and Marne, 1988; Mladineo, 1991; Pintér, 1996). All these algorithms require an evaluation of 2^m functions at each iteration, where m is the number of inputs, and hence they suffer from the curse of dimensionality. While such algorithms can still be implemented in an extremum seeking setting, the notation and derivation would be much more complicated and, hence, these details are omitted for simplicity.*

Note that $Q(\cdot)$ attains a maximum [4] on the compact interval $[a,b]$ since it is Lipschitz and, therefore, continuous. The maximum value of $Q(\cdot)$ is denoted as

$$y^* := \max_{u \in [a,b]} Q(u) \tag{3}$$

and the set of all u for which the global maximum is attained is defined as

$$\Phi := \{u \in [a,b] : Q(u) = y^*\}\ . \tag{4}$$

The Shubert algorithm generates a sequence of input points $u_1, u_2, ...$ within a closed interval $[a,b]$ by using measurements $y_1, y_2, ...$ as follows.

Shubert algorithm:

- Arbitrarily choose [5] an initial input $u_1 \in [a,b]$
- Find the next u_{k+1} such that the following equation is satisfied:

$$F_k(u_{k+1}) = M_k, \quad k = 1, 2, ..., \tag{5}$$

where

$$F_k(u) := \min_{j=1,\ldots,k} \{y_j + L|u - u_j|\}, \tag{6}$$

$$M_k := \max_{u \in [a,b]} F_k(u). \tag{7}$$

The sequence of functions $F_k(u)$ and numbers M_k for $k = 1, 2, \ldots$ possess properties that play a key role in the convergence analysis of the Shubert algorithm, and lead to the following theorem taken from (Shubert, 1972):

[3] The system is considered for $k = 1, 2, \ldots$ instead of $k = 0, 1, 2, \ldots$ as done in (Shubert, 1972); this is done to have consistent notation with the next section where the former index set more naturally arises when considering dynamic plants.

[4] Without loss of generality we concentrate on finding the maximum of $Q(\cdot)$; indeed, finding a minimum of a function $\widetilde{Q}(\cdot)$ can be done by defining $Q(\cdot) := -\widetilde{Q}(\cdot)$ and then finding a maximum of $Q(\cdot)$.

[5] Usually, the initial point is selected as $u_1 = \frac{a+b}{2}$.

Preprint submitted to Automatica
Received October 16, 2012 13:44:54 PST

Theorem 1 *Suppose that Assumption 1 holds and that the input sequence for system (1) is generated by the Shubert algorithm. Then the following holds*

$$\lim_{k\to\infty} y_k = y^* \ . \qquad \square$$

Remark 3 *It was shown in (Shubert, 1972) that a good estimate of the maximum y^* is given by:*

$$\hat{y}_k^* := \max_{j=1,\ldots,k} \{y_j\} \ . \tag{8}$$

Hence, Theorem 1 implies that $\hat{y}_k^*$ also converges to y^*. Moreover, the rate of convergence of $y^* - \hat{y}_k^*$ is of order $O\left(\frac{1}{k}\right)$ for all Lipschitz functions satisfying (2).

Remark 4 *Note that the main result in (Shubert, 1972) also shows that $\lim_{k\to\infty} M_k = y^*$ and $\lim_{k\to\infty} |u_k|_\Phi = 0$, where $|u|_\Phi := \inf_{z\in\Phi} |z-u|$ denotes the distance of the point u from the set Φ. For simplicity, we only state the convergence properties of y_k in Theorem 1 and in our main result (Theorem 2) which ensure an appropriate convergence to the global maximum y^*.*

Remark 5 *Note that the Shubert algorithm does not require derivatives of the map $Q(\cdot)$; in fact, $Q(\cdot)$ does not have to be differentiable everywhere*[6]*. Moreover, the algorithm finds* global extremum *on the compact set $[a, b]$ in presence of local extrema.*

3 Extremum seeking adaptation of Shubert algorithm

The diagram of the closed loop system with the proposed Shubert-based extremum seeking algorithm is shown in Figure 1. The control objective of the extremum seeking control is to drive the trajectories of the closed loop system in Figure 1 to eventually converge to a neighbourhood of the optimum y^* without the precise knowledge of the system model. Since the plant is dynamic, we will manage to achieve this only from a compact set of the plant initial states, i.e. the domain of attraction. Moreover, by increasing the sampling period T we will show that the domain of attraction of the closed loop can be made arbitrarily large and the neighboorhood of the optimum y^* to which the solutions converge can be made arbitrarily small; in other words, we achieve semi-global practical convergence.

The Shubert algorithm is applied to this dynamical system but since the algorithm from the previous section is discrete-time and the plant (9) is continuous-time, we need to use certain analog-to-digital and digital-to-analog converters.

[6] As the function is Lipschitz, it is differentiable almost everywhere according to Rademacher's Theorem (Clarke *et al.*, 1998).

Single-input single-output (SISO) dynamic plants are considered

$$\dot{x} = f(x, u), \quad x(0) = x_0, \tag{9}$$
$$y = h(x), \tag{10}$$

where $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}$ are locally Lipschitz in each argument. Assume that $u \in [a, b]$.

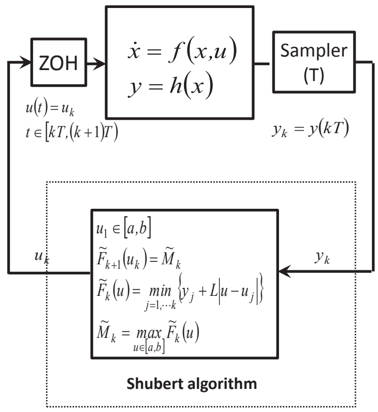

Fig. 1. The diagram of the closed loop system

It is assumed that equidistant sampling time instants $t_k := kT$ are given, where $T > 0$ is a design parameter that will be determined later. T is referred to as the *waiting time* for reasons that will be explained in the sequel; in our implementation it plays the role of the classical sampling period. It is assumed that the input to the plant is piecewise constant, that is we have:

$$u(t) = u(t_k) =: u_{k+1}, \qquad \forall t \in [t_k, t_{k+1}) \ , \tag{11}$$

for $k = 0, 1, \ldots$. Note that we use a slight abuse of notation where u_1 is applied on the time interval $[t_0, t_1)$; in particular, the index of the input sequence u_k takes values $k = 1, 2, \ldots$. To formally state our system we introduce some notation. For any constant input $u \in [a, b]$, the solution of the plant at time t and starting from an initial state x_0 is denoted $x(t, x_0, u)$; obviously $x(0, x_0, u_1) = x_0$. The notation $x_k := x(t_k, x_{k-1}, u_k)$ for $k = 1, 2, \ldots$ is used. In particular, the sampler collects the output measurements:

$$y_k := h(x_k) \qquad k = 1, 2, \ldots \ . \tag{12}$$

Note that with our notation we have that x_k and y_k can be defined for $k = 0, 1, \ldots$, whereas u_k only for $k = 1, 2, \ldots$. However, in our algorithm and analysis only y_k and u_k for $k = 1, 2, \ldots$ are needed. Hence, our abuse of notation.

Preprint submitted to Automatica
Received October 16, 2012 13:44:54 PST

Now we can apply Shubert algorithm from the previous section to the dynamic plant (9), (10) with sampler (12) and zero order hold (11).

Extremum seeking (Shubert) algorithm:

- Arbitrarily choose an initial input $u_1 \in [a, b]$
- Find the next u_{k+1} such that the following equation is satisfied:

$$\widetilde{F}_k(u_{k+1}) = \widetilde{M}_k, \quad k = 1, 2, ..., \tag{13}$$

where

$$\widetilde{F}_k(u) := \min_{j=1,\dots,k} \{y_j + L|u - u_j|\}, \tag{14}$$

$$\widetilde{M}_k := \max_{u\in[a,b]} \widetilde{F}_k(u). \tag{15}$$

Remark 6 *Note that in (14), (15), we use a different notation $\widetilde{F}_k(u)$ and $\widetilde{M}_k$ as opposed to $F_k(u)$ and M_k used in (6) and (7). This is because in this section y_j is the sampled output measurement of a dynamic system.*

The closed loop system consists of the plant (9), (10), the above algorithm and the sampler and zero order hold that are described above. The goal is to show that this closed loop system would achieve global extremum seeking under certain conditions. In order to prove our main results, we will assume that for each constant $u \in [a, b]$ there exists an equilibrium for (9) that is globally asymptotically stable. More precisely, we use the following:

Assumption 2 *There exists a locally Lipschitz function $\ell : [a, b] \to \mathbb{R}^n$ such that*

$$f(\ell(u), u) = 0, \qquad \forall u \in [a, b]. \tag{16}$$

Moreover, there exists $\beta \in \mathcal{KL}$ such that for any $u \in [a, b]$ and $x_0 \in \mathbb{R}^n$, the following inequality holds[7]

$$|x(t, x_0, u) - \ell(u)| \le \beta(|x_0 - \ell(u)|, t) \qquad \forall t \ge 0\ . \tag{17}$$

Remark 7 *Denote*

$$Q(\cdot) := h \circ \ell(\cdot) \tag{18}$$

as the steady-state input-to-output map of (9), (10). The same assumptions for $Q(\cdot)$ in (18) are used as in the previous section. Note that the domain of $Q(\cdot)$ is a compact interval $[a, b]$ and, hence, the function $Q(\cdot)$ achieves a maximum on the interval $[a, b]$. Moreover, note that $Q(\cdot)$ is locally Lipschitz since $h(\cdot)$ and $\ell(\cdot)$ are assumed to be locally Lipschitz. Hence, the first part of Assumption 1 holds for $Q(\cdot)$ given by (18). However, as we also need to know the Lipschitz constant L of $Q(\cdot)$, we will still explicitly state that Assumption 1 holds for $Q(\cdot)$ in (18).

[7] A function $\beta : \mathbb{R}_{\ge 0} \times \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ is of class $\mathcal{KL}$ if for each fixed $t \ge 0$ the function $\beta(\cdot, t)$ is continuous, zero at zero and strictly increasing and for each $s \ge 0$ the function $\beta(s, \cdot)$ is strictly decreasing to zero.

Remark 8 *It is easy to extend our results to infinite dimensional systems in a manner similar to (Teel and Popović, 2001). In this case, we would need to appropriately generalize Assumption 2. This level of generality is not pursued in order to keep the presentation simpler.*

Our goal is to show that the algorithm described above can find approximately the global maximum of $Q(\cdot)$ from an arbitrary set of initial conditions and to within an arbitrary prescribed margin if the waiting time T is sufficiently large.

4 Main result

This section contains the main result of this paper (Theorem 2) that shows appropriate semi-global practical convergence of the trajectories of the closed loop system from the previous section. The domain of convergence can be enlarged and the accuracy of the algorithm improved by increasing the waiting time T. To this end, we first state an auxiliary result that will allow us to carry out our analysis in discrete-time; its proof is given in the appendix.

Proposition 1 *Consider the closed loop system consisting of the plant (9), (10), sampler (12), zero order hold (11) and the extremum seeking algorithm. Suppose that Assumption 2 holds. Then, for any strictly positive pair (Δ, ν) there exists $T > 0$ such that for any $|x_0| \le \Delta$ and any $u_k \in [a, b]$, $k = 1, \dots$ we have that*

- $|x_k| \le \ell_{max} + 1$ *for* $k = 1, 2, \dots$
- $|y_k - Q(u_k)| \le \nu$ *for all* $k = 1, 2 \dots,$

where $\ell_{max} := \max_{u\in[a,b]} |\ell(u)|$. □

Remark 9 *Note that we can not claim in Proposition 1 that $|y_0 - \ell(u_1)| \le \nu$; indeed, the proof uses the stability properties stated in Assumption 2 to bound the solutions after the waiting time T has elapsed. Hence the name "waiting time".*

The main result of this paper is stated next:

Theorem 2 *Consider the closed loop system consisting of the plant (9), (10), sampler (12), zero order hold (11) and the extremum seeking algorithm. Suppose Assumptions 1 and 2 hold. Then, for any strictly positive (Δ, ν), there exists $T > 0$ such that for any $|x_0| \le \Delta$ we have that the following holds for the closed loop system:*

$$y^* - \nu \le \liminf_{k\to\infty} y_k \le \limsup_{k\to\infty} y_k \le y^* + \nu\ . \qquad \square$$

Preprint submitted to Automatica
Received October 16, 2012 13:44:54 PST

Remark 10 *The convergence rate of the sampled outputs* $y^* - y_k$ *is* $O\left(\frac{1}{k}\right)$*, see Remark 3, but the rate of convergence of* $y(t)$ *to the* ν*-neighborhood of the origin is of the order* $T \cdot O\left(\frac{1}{k}\right)$*, where* T *is the waiting time. Note that Theorem 2 establishes semi-global practical convergence of the closed loop in the parameter* T*. In particular, the domain of attraction* $|x_0| \leq \Delta$ *can be arbitrarily large and the accuracy of the algorithm that is measured by* ν *can be arbitrarily small. However, the waiting time* T *is necessarily larger for larger domain of attraction (larger* Δ*) and/or better accuracy of the algorithm (smaller* ν*) and, hence, the convergence of* $y(t)$ *to a neighborhood of the maximum* y^* *is slower. This tradeoff was observed in other, e.g. gradient based, extremum seeking schemes, see (Tan* et al.*, 2006).*

Remark 11 *Note that the domain* $[a, b]$ *of the map* $Q(\cdot)$ *that we are optimizing is compact. Hence, global optimization is with respect to this compact interval.*

Remark 12 *The proof that we present below is an appropriate generalization of the proof in (Shubert, 1972). In particular, we will show below that the sampled behaviour of the dynamic closed loop system given by (22) is a perturbed version of the algorithm for static plants given by (1). More precisely, we show that the algorithm (1) is robust to small additive perturbations. In particular, formally if we let* $\nu = 0$ *in conclusion of Theorem 2 we get that*

$$y^* \leq \liminf_{k\to\infty} y_k \leq \limsup_{k\to\infty} y_k \leq y^* \ ,$$

which implies that

$$\liminf_{k\to\infty} y_k = \limsup_{k\to\infty} y_k = \lim_{k\to\infty} y_k = y^*.$$

Hence, we can recover the result of Theorem 1 if we consider static plants. However, we note that for dynamic plants with finite waiting times it is not reasonable to expect that $\nu = 0$ *in Proposition 1; consequently, Theorem 2 is stated with strictly positive* ν*.*

Remark 13 *Our results can be easily extended to the case when the measured output is contaminated by measurement noise, which is a more realistic scenario. We note that all our results can be appropriately adjusted in the following manner. Suppose that the measured output is* $y(t) = h(x(t)) + n(t)$*, where* $n(t)$ *is the noise satisfying*

$$ess\ sup_{t\in[0,\infty)} |n(t)| \leq \nu_m \qquad (19)$$

for some $\nu_m > 0$*. Theorem 2 can be rephrased as follows.*

Remark 14 *Suppose Assumptions 1, 2 and (19) hold. Then, for any strictly positive* (Δ, ν)*, there exist* $T > 0$ *and* $\nu_m > 0$ *such that for any* $|x_0| \leq \Delta$ *we have that the following holds for the closed loop system:*

$$y^* - \nu \leq \liminf_{k\to\infty} y_k \leq \limsup_{k\to\infty} y_k \leq y^* + \nu \ . \qquad (20)$$

The proof of the above fact follows almost the same steps as the proof of Theorem 2 and it is omitted for simplicity.

Remark 15 *Our result is somewhat similar to results in (Teel and Popović, 2001) where a framework for extremum seeking control design method was presented for nonlinear programming (NLP) type optimization algorithms. However, the proof of our result is different from that in (Teel and Popović, 2001).*

(1) Note that while results in (Teel and Popović, 2001) were motivated by NLP algorithms, they are general enough to apply to any algorithm of the type

$$u^+ \in F(u, G(u)) \qquad (21)$$

that satisfies appropriate Lyapunov conditions for uniform asymptotic stability of the set of minimizers Φ*; note that in principle sampling algorithms, like the Shubert algorithm, could be used within the framework of (Teel and Popović, 2001) if they satisfy appropriate conditions. Indeed, Theorem 1 states that the Shubert algorithm yields attractivity of* Φ*. Our main result (Theorem 2) states that attractivity of* Φ *is robust under perturbations induced the plant dynamics. Note that Shubert algorithm is not uniformly stable, which is required to use results in (Teel and Popović, 2001). Indeed, if we initialize Shubert algorithm from the set of minimizers* Φ*, the algorithm would generate trajectories that exit* Φ *and then take time to converge to it; this behavior is not possible for uniformly stable systems which would have the set* Φ *as its set of equilibria.*

(2) Note that in (Teel and Popović, 2001), the stability properties NLP algorithm were shown for any trajectory generated from the NLP algorithm. In the Shubert algorithm, the initial condition of input sequence is usually fixed. Our result (Theorem 2) shows the robustness of the Shubert algorithm with this given initial condition.

(3) We have used an alternative proof technique to (Teel and Popović, 2001) in order to prove our main results; a subtle difference with (Teel and Popović, 2001) is that Theorem 2 states (practical) convergence from a compact set $u_0 \in [a, b]$ *whereas results in (Teel and Popović, 2001) assume the domain of attraction of (21) to be* $\mathbb{R}^n$ *and conclude (practical) convergence from compact subsets* $u_0 \in D \subset \mathbb{R}^n$*.*

We note that the Shubert algorithm may produce control inputs u_k that yield undesirable plant transients for some plants since they may be "too aggressive" with large jumps at different consecutive samples. Modifying

Preprint submitted to Automatica
Received October 16, 2012 13:44:54 PST

the algorithm to use "less aggressive" sampling techniques is an interesting topic for further research; however, we believe that this issue is outside the scope of this paper. We present some simulation results in the sequel to illustrate what the sequence u_k may look like in an example.

Proof of Theorem 2: Let (Δ, ν) be given and let $\left(\Delta, \frac{\nu}{3}\right)$ generate $T > 0$ via Proposition 1. Consider an arbitrary $|x_0| \leq \Delta$ and the corresponding sequence of measurements y_k and control inputs u_k that result from the closed loop system. Note that for all $k = 1, 2, \ldots$, we can write

$$y_k = Q(u_k) + w_k \ , \tag{22}$$

where $w_k := y_k - Q(u_k)$ and $Q(\cdot)$ comes from (18). The controller in the closed loop system produces a corresponding sequence of inputs u_k and this results in a sequence of outputs y_k. Note that for every different initial condition x_0 we obtain different sequences u_k, y_k, w_k; however, since we fixed x_0, all sequences are fixed. Moreover, from Proposition 1 we always have that $|w_k| \leq \frac{\nu}{3}$ holds for all $k = 1, 2, \ldots$.

The first observation is that since we have that $y_k = Q(u_k) + w_k \leq y^* + \frac{\nu}{3} \leq y^* + \nu$ for all k, then we have that

$$\limsup_{k\to\infty} y_k \leq y^* + \nu \ . \tag{23}$$

Hence, the last inequality in Theorem 2 holds. Moreover, by definition we have that $\liminf_{k\to\infty} y_k \leq \limsup_{k\to\infty} y_k$ and the only thing left to prove is that

$$\liminf_{k\to\infty} y_k \geq y^* - \nu \ . \tag{24}$$

Note that since $\widetilde{M}_k$ is a nonincreasing sequence that is bounded, it has a limit that is denoted as $\widetilde{M} = \lim_{k\to\infty} \widetilde{M}_k$. Obviously, we have that

$$\widetilde{M}_k \geq \widetilde{M}, \qquad \forall k \ . \tag{25}$$

Let z be a limit point of u_k; that is, there exists a subsequence u_{n_k} of the sequence u_k that converges to z. Note that if u_k is bounded, by Bolzano-Weierstrass Theorem (Fitzpatrick, 2006), there exists at least one limit point. Moreover, denote as $\mathcal{Z} \subset \mathbb{R}$ the set of all limit points of u_k. Then, since $Q(\cdot)$ is continuous we have that

$$\liminf_{k\to\infty} Q(u_k) = \inf_{z\in\mathcal{Z}} Q(z) \ . \tag{26}$$

Consider the set $\mathcal{U} := \{u \in [a, b] : u = u_k \text{ for some k} = 1, 2, \ldots\}$ in two different cases.

Case 1: The set $\mathcal{U}$ has infinitely many distinct elements. There is no loss of generality to assume that $u_m \neq u_n$ for all $m \neq n$. In this case, the set $\mathcal{U}$ is infinite and since it is bounded it has at least one limit point z. It is shown next that for any limit point $z \in \mathcal{Z}$ we have that the following holds:

$$Q(z) \geq \widetilde{M} - \frac{\nu}{3} \ . \tag{27}$$

For the purpose of showing contradiction, assume that there exists some arbitrary (small) $\epsilon > 0$ so that

$$Q(z) \leq \widetilde{M} - \frac{\nu}{3} - \epsilon \ . \tag{28}$$

Let u_{k_n} be the sequence of points converging to z and let $n(\epsilon)$ be such that

$$n \geq n(\epsilon) \ \Rightarrow \ |u_{k_n} - z| < \frac{\epsilon}{2L} \ . \tag{29}$$

Using (2), (28) and (29), we conclude that $n \geq n(\epsilon)$ implies:

$$Q(u_{k_n}) + w_{k_n} \leq L\,|u_{k_n} - z| + Q(z) + \frac{\nu}{3} < \widetilde{M} - \frac{\epsilon}{2} \tag{30}$$

Moreover, from (14) we have that for all $k \geq k_{n(\epsilon)}$ the following holds for all $u \in [a, b]$

$$\widetilde{F}_k(u) \leq \widetilde{F}_{k_n}(u) \leq L\,|u - u_{k_n}| + Q(u_{k_n}) + w_{k_n} \ . \tag{31}$$

Using (30) and (31) we conclude that for any $n \geq n(\epsilon)$ we have that $k \geq k_{n(\epsilon)}$ and $|u - u_{k_n}| \leq \frac{\epsilon}{2L}$ imply that

$$\widetilde{F}_k(u) \leq L\,|u - u_{k_n}| + Q(u_{k_n}) + w_{k_n} < \widetilde{M} \ . \tag{32}$$

This together with (25) and (13) implies that if $k \geq k_n$ then there does not exist $u_k \in \left[u_{k_n} - \frac{\epsilon}{2L}, u_{k_n} - \frac{\epsilon}{2L}\right]$ which contradicts the fact that z is a limit point. Hence, we conclude that all limit points $z \in \mathcal{Z}$ satisfy (27); hence, $\inf_{z\in\mathcal{Z}} Q(z) \geq \widetilde{M} - \frac{\nu}{3}$. Using the fact that $y_k \geq Q(u_k) - \frac{\nu}{3}$ for all k we can write using (27) that:

$$\begin{aligned}\liminf_{k\to\infty} y_k &\geq \lim \inf_{k\to\infty} Q(u_k) - \frac{\nu}{3} \\ &= \inf_{z\in\mathcal{Z}} Q(z) - \frac{\nu}{3} \geq \widetilde{M} - \frac{2\nu}{3} \ .\end{aligned} \tag{33}$$

Moreover, note that we also have for any $u^* \in \Phi$ that:

$$\begin{aligned}\widetilde{M}_k &= \max_{u\in[a,b]} \min_{j=1,\ldots,k} \{y_j + L|u - u_j|\} \\ &\geq \min_{j=1,\ldots,k} \{Q(u_j) + w_j + L|u^* - u_j|\} \\ &\geq Q(u^*) + \min_{j=1,\ldots,k} w_j \geq y^* - \frac{\nu}{3}, \qquad \forall k.\end{aligned} \tag{34}$$

Preprint submitted to Automatica
Received October 16, 2012 13:44:54 PST

Hence, the following holds

$$\widetilde{M} \geq y^* - \frac{\nu}{3} \tag{35}$$

and together with (33) the expression (24) holds in this case.

Case 2: The set $\mathcal{U}$ has finitely many distinct elements.

In this case, it is clear that there exists $n \geq 0$ such that $u_m = u_n$ for all $m > n$. The proof goes along similar lines as in the previous case but now we consider the sequence $u_m = u_n$ for all $m \geq n$. It is first proved that

$$Q(u_m) \geq \widetilde{M} - \frac{\nu}{3} \qquad \forall m \geq n \; . \tag{36}$$

For the purpose of showing contradiction we assume that there exists an arbitrarily small $\epsilon > 0$ such that

$$Q(u_m) \leq \widetilde{M} - \frac{\nu}{3} - \epsilon \qquad \forall m \geq n \; . \tag{37}$$

Consider an arbitrary $m \geq n$ and using (37) we conclude that

$$Q(u_m) + w_m \leq Q(u_m) + \frac{\nu}{3} \leq \widetilde{M} - \epsilon < \widetilde{M} \; , \tag{38}$$

but this implies that

$$\begin{aligned} \widetilde{F}_m(u_{m+1}) &= \widetilde{F}_m(u_m) \leq \widetilde{F}_n(u_m) \\ &\leq L\,|u_m - u_n| + Q(u_n) + w_n \\ &= Q(u_n) + w_n < \widetilde{M} \; , \end{aligned} \tag{39}$$

which contradicts the existence of u_{m+1} satisfying (13). Hence, we have that (36) holds and then this together with (35) implies that for all $m \geq n$ we have

$$\begin{aligned} y_m &= Q(u_m) + w_m \geq Q(u_m) - \frac{\nu}{3} \\ &\geq \widetilde{M} - \frac{2\nu}{3} \geq y^* - \nu \; . \end{aligned} \tag{40}$$

Hence, (24) holds and this completes the proof.

5 A Numerical Example

In order to illustrate the effectiveness of the proposed global extremum seeking algorithm, a simple linear-time-invariant system is considered.

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \begin{bmatrix} -2 & 1 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, x_0 = \begin{bmatrix} 5 \\ 2 \end{bmatrix} \\ y &= x_1 - \sin(3x_1) + 1. \end{aligned} \tag{41}$$

A simple calculation yields $\ell(u) = -(A)^{-1}Bu$, where $A := \begin{bmatrix} -2 & 1 \\ 0 & -3 \end{bmatrix}$ and $B := \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. The input-output map of the system becomes

$$Q(u) = \frac{u}{6} - \sin\left(\frac{u}{2}\right) + 1.$$

As shown in Figure 2, $Q(\cdot)$ has 3 maxima with a global maximum $y^* = 7.82$ (red dotted line in Figure 3 and Figure 5) at $u^* = 35.24$ when the input u is in a compact set $[0, 39]$. This $Q(\cdot)$ comes from (Hansen *et al.*, 1992*b*, Problem 19, Table 1).

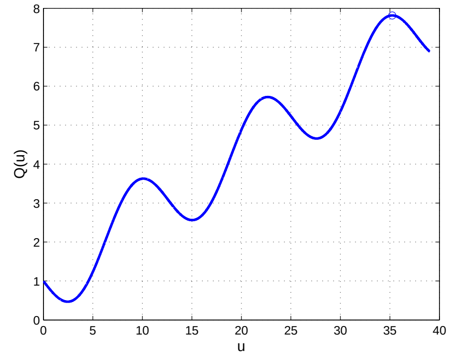

Fig. 2. The input-to-output mapping

First, we fix the Lipschitz constant $L = \frac{2}{3}$ which comes from (Hansen *et al.*, 1992*b*) and compare the performance of global extremum seeking for two different choices of the waiting time. By using MATLAB, the settling time (2%) of this system can be obtained as $2.5s$[8], $T = 4$ (larger than the settling time) and $T = 0.1$ (much smaller than the settling time) are used.

Figure 3 clearly shows that the longer waiting time T, the better accuracy (smaller ν) will be. The longer waiting time also leads to a slow convergence. There is an obvious design trade-off between the accuracy and the convergence speed.

Figure 4 also shows first 49 inputs obtained from the input sequence computed from Shubert algorithm for dynamic systems when $T = 4$. It indicates that the input sequence converges quickly to a small neighborhood of the optimal input $u^* = 35.24$.

Next, we investigate the performance of the global extremum seeking when a conservative estimate of the Lipschitz constant is used, for example $L = 5$ (instead of $\frac{2}{3}$). For simplicity of presentation, we fix $T = 4$ to compare the effect of estimated L. From Figure 5, it clearly shows

[8] Usually, the settling time is a good candidate for the waiting time.

Preprint submitted to Automatica
Received October 16, 2012 13:44:54 PST

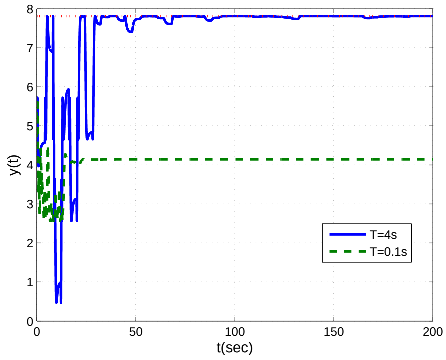

Fig. 3. The output for $x_0 = (5\ 2)^T$, $L = \frac{2}{3}$ and $T = 4$s or $T = 0.1$s.

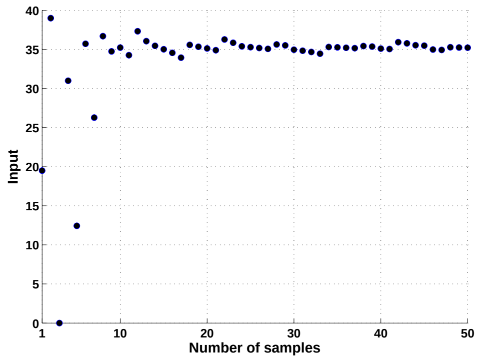

Fig. 4. Evolution of the input: $L = \frac{2}{3}$ and $T = 4$s

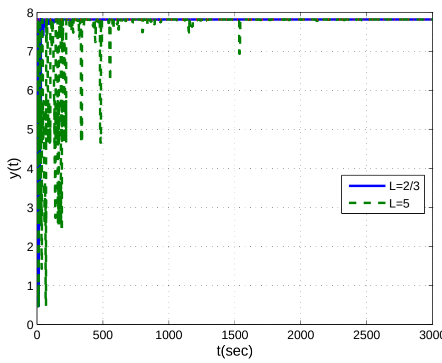

Fig. 5. The output for $x_0 = (5\ 2)^T$, $T = 4$s and $L = \frac{2}{3}$ or $L = 5$.

that when a conservative L is used, the proposed global extremum seeking still works. However, the convergence speed is much slower.

6 Conclusion

This note proposed a global extremum seeking scheme for nonlinear dynamic systems by combining a “sampling optimization method”, the so-called Shubert algorithm, and dynamic nonlinear plants. Global optimization is proved on compact sets of inputs for static plants and semi-global practical convergence is achieved for dynamic plants where the waiting time T is the parameter that needs to be adjusted. Numerical simulations are provided to demonstrate of the proposed scheme. This work opens new research opportunities for the adaptation of sampling optimization algorithms in (Strongin and Sergeyev, 2000) in the context of extremum seeking control.

References

Ariyur, K.B. and M. Krstić (2003). *Real-Time Optimization by Extremum Seeking Control.* Wiley-Interscience. Hoboken, NJ.

Clarke, F.H., Yu.S.Ledyaev, R.J. Stern and P.R. Wolenski (1998). *Nonsmooth analysis and control theory.* Springer Verlag. New York, Inc.

Fitzpatrick, Patrick M. (2006). *Advanced Calculus (2nd ed.).* Thompson Brooks/Cole. Belmont, CA.

Hansen, P., B. Jaumard and S.-H. Lu (1992*a*). Global optimization of univariate Lipschitz functions: I. Survey and properties. *Mathematical Programming* **55**(1), 251–272.

Hansen, P., B. Jaumard and S.-H. Lu (1992*b*). Global optimization of univariate Lipschitz functions: II. New algorithms and computational comparison. *Mathematical Programming* **55**(1), 273–292.

Hansen, P., B. Jaumard and S. H. Lu (1992*c*). On using estimates of lipschitz constant in globl optimization. *Journal of Optimization Theory and Applications* **75**, 195–200.

Meewella, C. C. and D. Q. Marne (1988). An algorithm for global optimization of lipschitz functions. *Journal of Optimization Theory and Applications* **57**, 307–322.

Mladineo, R. H. (1986). An algorithm for finding the global maximum of a multi- modal, multvariate function. *Mathematical Programming* **34**, 188–200.

Mladineo, R. H. (1991). *Stochastic minimization of Lipschitz functions, In Recent Advances in Global Optimization, Princeton Series in Computer Science.* Princeton University Press, Princeton.

Moase, W.H., C. Manzie and M.J. Brear (2010). Newton-like extremum-seeking for the control of thermoacoustic instability. *IEEE Transactions on Automatic Control* **55**(9), 2094–2105.

Pintér, J. (1996). *Global optimization in action, In volume 6 of Nonconvex Optimization and Its Applications.* Kluwer Academic Publishers, Dordrecht.

Shubert, B.O. (1972). A sequential method seeking the global maximum of a function. *SIAM Journal on Numerical Analysis* **9**(3), 379–388.

Strongin, R. G. (1973). On the convergence of an algorithm for finding a global extremum. *Engineering Cybernetics* **11**, 549–555.

Strongin, R.G. and Y.D. Sergeyev (2000). *Global Op-*

Preprint submitted to Automatica
Received October 16, 2012 13:44:54 PST

timization with Non-Convex Constraints - Sequential and Parallel Algorithms. Springer-Verlag. New York.

Tan, Y., D. Nešić and I.M.Y. Mareels (2006). On non-local stability properties of extremum seeking control. *Automatica* **42**(6), 889–903.

Teel, A.R. and D. Popović (2001). Solving smooth and nonsmooth multivariable extremum seeking problems by the methods of nonlinear programming. In: *Proceedings of the 2001 American Control Conference*. Vol. 3. pp. 2394–2399.

7 Appendix

Proof of Proposition 1: Let $\beta(\cdot,\cdot)$ and $\ell(\cdot)$ come from Assumption 2. Let (Δ,ν) be given. Let $L_h \geq 0$ be the Lipschitz constant of $h(\cdot)$ on the set $|x| \leq \ell_{max}+1$ and let $T > 0$ be such that

$$\beta\left(\Delta + 2\ell_{max} + 1, T\right) \leq \min\left\{\frac{\nu}{L_h(2\ell_{max}+1)}, 1\right\} .$$

Consider an arbitrary $|x_0| \leq \Delta$ and $u_k \in [a,b]$ for all k. Then, from our choice of T, Assumption 2 and the fact that $\beta \in \mathcal{KL}$ we have that

$$|x_1 - \ell(u_1)| \leq \beta(|x_0 - \ell(u_1)|, T) \leq \beta(\Delta + \ell_{max}, T) \leq 1 ,$$

which implies $|x_1| \leq \ell_{max}+1$; this, in turn implies $|x_1 - \ell(u_2)| \leq 2\ell_{max}+1$. Moreover, by using our choice of T, the time invariance of (9), $\beta \in \mathcal{KL}$ and induction we have that for all $k \geq 2$ the following holds

$$\begin{aligned} |x_k - \ell(u_k)| &\leq \beta(|x_{k-1} - \ell(u_k)|, T) \\ &\leq \beta(\Delta + 2\ell_{max} + 1, T) \\ &\leq 1 . \end{aligned} \tag{42}$$

Hence, we have that $|x_k| \leq \ell_{max}+1$ for all $k \geq 1$, which proves the first claim. Then, using our choice of T it directly follows that for all $k \geq 1$:

$$\begin{aligned} |y_k - Q(u_k)| &= |h(x_k) - h \circ \ell(u_k)| \\ &\leq L_h |x_k - \ell(u_k)| \\ &\leq L_h \beta(|x_{k-1} - \ell(u_k)|, T) \\ &\leq L_h \beta(\Delta + 2\ell_{max} + 1, T) \\ &\leq \nu , \end{aligned} \tag{43}$$

which completes the proof.

Preprint submitted to Automatica
Received October 16, 2012 13:44:54 PST