University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Khong, SZ;Nesic, D;Manzie, C;Tan, Y

Title:
Multidimensional global extremum seeking via the DIRECT optimisation algorithm

Date:
2013-07-01

Citation:
Khong, S. Z., Nesic, D., Manzie, C. & Tan, Y. (2013). Multidimensional global extremum seeking via the DIRECT optimisation algorithm. AUTOMATICA, 49 (7), pp.1970-1978. https://doi.org/10.1016/j.automatica.2013.04.006.

Persistent Link:
https://hdl.handle.net/11343/299556

CONFIDENTIAL. Limited circulation. For review only

Multidimensional global extremum seeking via the DIRECT optimisation algorithm ⋆

Sei Zhen Khong [a], Dragan Nešić [a], Chris Manzie [b], Ying Tan [a]

[a] *Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, VIC 3010, Australia*

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

Abstract

DIRECT is a sample-based global optimisation method for Lipschitz-continuous functions defined over compact multidimensional domains. This paper adapts the DIRECT method with a modified termination criterion for global extremum seeking control of multivariable dynamical plants. Finite-time semi-global practical convergence is established based on a periodic sampled-data control law, whose sampling period is a parameter which determines the region and accuracy of convergence. A crucial part of the development is dedicated to a robustness analysis of the DIRECT method against bounded additive perturbations on the objective function. Extremum seeking involving multiple units is also considered within the same context as a means to increase the speed of convergence. Numerical examples of global extremum seeking based on DIRECT are presented at the end.

Key words: Extremum seeking control, multidimensional global optimisation, DIRECT method, robustness analysis.

Nomenclature

$\mathbb{N}$	The natural numbers
$\mathbb{R}$	The real numbers
$\|\cdot\|_2$	The Euclidean norm on $\mathbb{R}^m$
$\|x\|_{\mathcal{Y}}$	The distance of a point x of a Banach space $\mathcal{X}$ with norm $\|\cdot\|$ from subset $\mathcal{Y} \subset \mathcal{X}$. That is, $\|x\|_{\mathcal{Y}} := \inf_{a\in\mathcal{Y}} \|a - x\|$
$\|Q\|_\infty$	The infinity norm of $Q : \Omega \to \mathbb{R}$, i.e. $\|Q\|_\infty := \sup_{u\in\Omega} \|Q(u)\|$

$\mathbb{Q}(Q, \nu) := \left\{ \tilde{Q} : \Omega \to \mathbb{R} \;\middle|\; \|\tilde{Q} - Q\|_\infty \le \nu \right\}.$

1 Introduction

Extremum seeking is a control scheme which optimises the steady-state input-output behaviour of a dynamical system, for which a precise mathematical model may *not* be readily available (Ariyur and Krstić, 2003; Zhang and Ordóñez, 2011). By prescribing certain generic stability, attractivity, and robustness properties of a class of continuous-time dynamical systems and a class of discrete-time nonlinear programming methods, Teel and Popović (2001) propose a unifying framework in which extremum seeking can be achieved with a sampled-data type controller. There, convergence of the method is established via Lyapunov-based arguments. The underlying idea of being able to apply a large class of optimisation algorithms in extremum seeking control inspired the recent paper by Nešić et al. (2010), which provides a similar unifying framework based on continuous-time optimisation methods whose convergence is proven using averaging and singular perturbation techniques described in Tan et al. (2006).

Despite the generality of the above frameworks, two downsides may be identified. First, they require online estimation of the derivatives of the steady-state input-output map, which is often inaccurate due to the presence of system dynamics. Certain regularity conditions are stipulated in Tan et al. (2006); Teel and Popović (2001) to guarantee robustness to such inaccuracy, but they may be difficult to verify. Second, a particular state-update form describing the optimisation algorithms (for e.g., gradient descent, Newton-Raphson (Boyd and Vandenberghe, 2004)) is stipulated, along with their asymptotic stability. Often outputs from algorithms of this

⋆ This work was supported by the Australian Research Council (DP0985388). This paper was not presented at any IFAC meeting. Corresponding author S. Z. Khong. Tel. +61 3 83440374. Fax +61 3 83447412.

Email addresses: szkhong@unimelb.edu.au (Sei Zhen Khong), dnesic@unimelb.edu.au (Dragan Nešić), manziec@unimelb.edu.au (Chris Manzie), yingt@unimelb.edu.au (Ying Tan).

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST

form tend to be entrapped at *local* extrema. Furthermore, certain optimisation methods cannot be written in a state-update form with the required stability property; see, for instance, Strongin and Sergeyev (2000).

Several sampling-based discrete-time Lipschitzian optimisation methods for a static map, which do not rely on derivatives in their formulation, can be found in the literature (Pintér, 1996). These deterministic methods are capable of locating a *global* extremum of a Lipschitz continuous function in the multiextremal case within a *compact* domain of search. The assumption that the system inputs lie within a compact set is physically well-motivated by the fact that saturation constraints are inherent in real-world actuators. The so-called Piyavskij-Shubert algorithm (Piyavskij, 1972; Shubert, 1972) has been adapted for extremum seeking control of dynamical plants in Nešić et al. (2012), where periodic sampled-data controllers are used along the lines of Teel and Popović (2001). It is worth noting that the Piyavskij-Shubert algorithm does not fit in the aforementioned frameworks of Nešić et al. (2010); Teel and Popović (2001). The extremum seeking convergence proof therein is carried out through careful robustness analysis of the specific algorithm itself. However, for general application to multivariate function optimisations the Piyavskij-Shubert algorithm suffers a serious drawback: an exponentially increasing running time leads to computational intractability.

Towards addressing the above issue, this paper considers another Lipschitzian optimisation method, the so-called DIRECT (DIviding RECTangles) (Jones et al., 1993). DIRECT searches through all 'possible' Lipschitz bounds on the objective function, and therefore operates intelligently between the local and global levels, the former of which is important in speeding up the rate of convergence once a neighbourhood around a global extremum point is found. The search space of both the Piyavskij-Shubert and DIRECT methods is composed of rectangles. By contrast with sampling all the vertices as in Piyavskij-Shubert, DIRECT only requires the sample of a rectangle's midpoint, irrespective of the dimensionality. As such, global optimisation of functions of several variables is computationally feasible with DIRECT. The convergence of DIRECT, however, relies on the fact that the input samples form a dense subset of the domain of search. This property is undesirable in extremum seeking control because even after DIRECT locates a neighbourhood of a global extremum at any point of time, there always exists a future time at which DIRECT samples outside this basin, which amounts to leaving the neighbourhood of the global extremum point and probing the system elsewhere. This process does *not* eventually relinquish, as is necessary for dense sampling. Therefore, a sensible termination procedure for the algorithm, which guarantees a point within a certain neighbourhood of a global optimum has been sampled, is needed from an extremum seeking control perspective.

This paper adopts a *modified* DIRECT method for global extremum seeking control of asymptotically stable time-invariant dynamical systems of possibly *infinite* dimension with *multiple* inputs belonging to a compact set, in a similar spirit to that of Nešić et al. (2012). A periodic sampled-data control scheme is proposed to achieve semi-global practical convergence in finite time, after which DIRECT is terminated with the input resulting in the closest sample to a global extremum selected onwards as a constant reference to the system. Tuning guidelines are provided for two design parameters, namely the sampling/waiting period and the accuracy of DIRECT's estimate before termination. Knowledge of the Lipschitz constant of the steady-state input-output map is assumed and exploited in deciding on when to terminate DIRECT. The unmodified DIRECT method for a static mapping requires no knowledge of its Lipschitz constant to work.

An important part of the convergence proof extends that of the original DIRECT method for a static objective function to one which is perturbed by the dynamics of the plant, in which bounds on the perturbations' magnitude at sampling instants can be controlled by the waiting time. Through the sampled-data devices, DIRECT optimises the steady-state behaviour of the dynamical system in the presence of these perturbations. Indeed, the waiting time is equal to the sampling period of the control law. These two terms are used interchangeably throughout the paper.

When multiple plants are available, computational parallelism may be exploited to speed up the convergence to a global extremum; see Esmaeil-Zadeh-Azar (2010); Strongin and Sergeyev (2000); Woodward et al. (2009). The aforementioned extremum seeking control scheme is also generalised in this paper to multi-unit extremum seeking problems without generating redundancy in sampling points. Sufficient conditions which ensure finite-time semi-global practical convergence are provided.

The paper has the following structure. The next section contains a recapitulation of the DIRECT method from Jones et al. (1993). Analysis of the robustness of DIRECT to small additive perturbations on the objective function is performed in Section 3. Subsequently in Section 4, DIRECT is adapted into extremum seeking of multi-input single-output (MISO) dynamic systems. Multi-unit extremum seeking is considered in Section 5. To illustrate the results, simulation examples are provided in Section 6.

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST

2 The DIRECT method

Consider the following bound-constrained optimisation problem:

$$y^* := \min_{u \in \Omega} Q(u), \tag{1}$$

where

$$\Omega := \{u \in \mathbb{R}^m \mid u_i \in [a_i, b_i] \subset \mathbb{R}, i = 1, 2, \ldots, m\} \tag{2}$$

and the following assumption holds.

Assumption 1 *$Q : \Omega \subset \mathbb{R}^m \to \mathbb{R}$ is a globally Lipschitz continuous function (Khalil, 2002), i.e. there exists a known $L > 0$ such that*

$$|Q(u) - Q(u')| \leq L\|u - u'\|_2$$

for all $u, u' \in \Omega$.

Assume without loss of generality that $a_i = 0$ and $b_i = 1$ for all $i = 1, \ldots m$, i.e. Ω is a unit hypercube, which can be obtained by appropriately normalising (1). Note that Q must possess a global minimum by the Extreme Value Theorem (Rudin, 1976, Thm. 4.16) since it is continuous on a compact domain. The map Q can be viewed as a MISO *static* system (i.e. without dynamics) and DIRECT (Jones et al., 1993) is a deterministic sampling method which solves (1). The only assumption that DIRECT makes is the Lipschitz continuity of Q; no knowledge of the system model is needed. This section gives a brief review of the DIRECT optimisation method (Jones et al., 1993) and Section 4 demonstrates how it can be applied for extremum seeking control of MISO *dynamical* systems.

The following describes the DIRECT algorithm; see Jones et al. (1993) for more details, including a pictorial illustration. Also see Finkel (2004) for a numerical implementation of the algorithm using MATLAB.

Algorithm 1 *The DIRECT algorithm.*

Given: A Lipschitz function $Q : \Omega \subset \mathbb{R}^m \to \mathbb{R}$.
Notation: q denotes the iteration number of DIRECT and k the total number of samples. Input samples taken by DIRECT are denoted $u_i \in \Omega$, $i = 0, 1, \ldots, k-1$ and the corresponding outputs $y_{i+1} := Q(u_i)$.

(i) *Initialise $q := 1$ and $k := 0$.*
(ii) *Evaluate $Q(u_0)$, where $u_0 \in \mathbb{R}^m$ denotes the centre point of Ω. Set $\hat{y}_1 := Q(u_0)$ and increment k, i.e. $k^+ := k + 1$.*
(iii) *Identify the set of indices of potentially optimal hyper-rectangles S, i.e. all $j \in \{0, \ldots, k-1\}$ for which there exists a positive $\tilde{L} \in \mathbb{R}$ such that*

$$\begin{aligned} Q(u_j) - \tilde{L}d_j &\leq Q(u_i) - \tilde{L}d_i \; \forall i = 0, \ldots, k-1 \\ Q(u_j) - \tilde{L}d_j &\leq \hat{y}_k - \epsilon|\hat{y}_k| \text{ for some } \epsilon > 0, \end{aligned} \tag{3}$$

where d_i denotes the distance from the centre point to the vertices of the i^{th} hyper-rectangle. This step can be performed efficiently using the Graham's scan (Graham, 1972), which is an algorithm for determining the convex hull of a finite set; see Jones et al. (1993).
(iv) *For every $j \in S$, subdivide the j^{th} hyper-rectangle with centre u_j according to the following rule. First identify the set I of dimensions $i \in \{1, \ldots, m\}$ in which the j^{th} hyper-rectangle has the maximum side length and let δ be one-third of this value. Sample Q at the points $u_j \pm \delta e_i$ for all $i \in I$, where e_i denotes the i^{th} unit vector in $\mathbb{R}^m$. Sequentially divide the j^{th} hyper-rectangle into thirds along the dimension $i \in I$ in an ascending order of $\min\{Q(u_j + \delta e_i), Q(u_j - \delta e_i)\}$. Set $k^+ := k + \Delta k$, where Δk is the number of* new *points sampled during the q^{th} iteration.*
(v) *Set $q^+ := q + 1$ and the minimal estimate*

$$\hat{y}_q := \min_{i=1,\ldots k} y_i. \tag{4}$$

(vi) *Loop from* (iii).

Remark 2 *The ϵ in (3) is a balance parameter which serves to prevent DIRECT from entrenching itself in the local search. Robustness of the algorithm to this parameter is analysed in detail in Finkel (2005). To alleviate the sensitivity to this parameter, which when inappropriately set may compromise the efficiency of DIRECT for certain problems, a modification with a time-varying ϵ is also proposed therein.*

Remark 3 *Since DIRECT identifies all potentially optimal hyper-rectangles via (3) for subdivision, it is well-balanced between local and global searches. Indeed, once DIRECT locates the basin of convergence of a global optimum, the local part of the algorithm automatically exploits it to accelerate the search (Jones et al., 1993). A modification of DIRECT which is strongly biased towards the local search can be found in Gablonsky (2001); Gablonsky and Kelley (2001).*

Proposition 4 (Jones et al. (1993)) *As the number of iterations approaches infinity, the points sampled by DIRECT form a dense subset of Ω. Because Q is Lipschitz continuous and hence continuous, this implies that the estimate by DIRECT converges to y^* in (1). Mathematically, it holds that*

$$\lim_{q \to \infty} \hat{y}_q = y^* := \min_{u \in \Omega} Q(u).$$

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST

The denseness in domain-sampling property of DIRECT mentioned in the proposition above is critical in the proof of its convergence. It holds by the way the algorithm is set up, which subdivides all potentially optimal rectangles in the search space iteratively. The Piyavskij-Shubert algorithm does not share this feature and is thus based on a different convergence analysis (Nešić et al., 2012; Piyavskij, 1972; Shubert, 1972).

Remark 5 *In the generalised extremum-seeking framework of Teel and Popović (2001), nonlinear optimisation algorithms are assumed to take a difference inclusion form:*

$$u_{k+1} \in F(u_k, G(u_k)), \tag{5}$$

where $u_k \in \Omega$ denotes the sample point at 'time' k, F is a 'state-update' set-valued map and G is a function that carries information regarding the estimate of the gradient of Q around u. Furthermore, (5) is required to satisfy a type of asymptotic stability property as guaranteed by the existence of a corresponding Lyapunov-like function. The DIRECT Algorithm 1 cannot be classified as being in this class of optimisation algorithms, since expressing it in the form of (5) does not seem possible and by virtue of the fact that the input samples u_k for $k = 1, 2, \ldots$ are generated in such a way that they form a dense subset of Ω, the asymptotic stability property stipulated in Teel and Popović (2001) does not hold.

3 Robustness analysis of DIRECT

3.1 Termination of DIRECT

The standard termination criterion of the DIRECT algorithm is a pre-specified number of iterations. Using knowledge of the Lipschitz constant L of $Q(\cdot)$, a lower bound on Q can be computed and progressively improved based on each new sample y_i. A less heuristic stopping criterion for DIRECT can then be y_i being within some tolerance of the bound, as is the case for Piyavskij-Shubert algorithm (Piyavskij, 1972; Shubert, 1972).

To be specific, suppose after a number of DIRECT iterations, there are J number of hyper-rectangular subblocks in the input space Ω. Consider the j^{th} rectangle with centre u_j, and let the distance between its vertices and u_j be d_j. Then by the Lipschitz continuity Assumption 1, the lowest possible value Q can attain on this rectangle is given by $Q(u_j) - \eta_j$ with $\eta_j := Ld_j$. It follows that

$$E := \min_{j=1,\ldots,J} (Q(u_j) - \eta_j) \leq \min_{u\in\Omega} Q(u), \tag{6}$$

i.e. E is an estimate of the lower bound on Q, which improves with each new sample DIRECT takes. Indeed, DIRECT can be programmed to terminate when a point within some η-tolerance from this lower bound is sampled, where $\eta := \max_j \eta_j$. The following lemma gives an estimate of the maximum number of iterations DIRECT needs to locate a point within a margin of η at a global minimum of Q.

Lemma 6 *Suppose Q in problem (1) satisfies the global Lipschitz condition in Assumption 1 with Lipschitz constant L. Given any $\eta > 0$, let*

$$N := 3^{m-1} \left(\frac{3^{m(i+1)} - 1}{3^m - 1} \right), \tag{7}$$

where m is the dimension of the input space as in (2) and $i \in \mathbb{N}$ is such that

$$\frac{(m3^{-2i})^{\frac{1}{2}}}{2} \leq \frac{\eta}{L}. \tag{8}$$

Then

$$\hat{y}_N - y^* \leq \eta,$$

where $\{\hat{y}_k\}_{k=1}^N$ is the sequence of estimates from applying DIRECT to any bounded function $\tilde{Q} : \Omega \subset \mathbb{R}^m \to \mathbb{R}$.

PROOF. The proof is established by combining aspects of Jones et al. (1993) and Gablonsky (2001). In particular, because hyper-rectangles in DIRECT's search space are constructed by dividing existing ones into thirds along selected dimensions in $\mathbb{R}^m$, the only possible side length a rectangle can have is 3^{-i} for $i = 0, 1, 2, \ldots$. Furthermore, since DIRECT always divides a rectangle on its largest side, it follows that after r divisions, a hyper-rectangle will have $j := r \mod m$ sides of length $3^{-(i+1)}$ and $m - j$ sides of length 3^{-i}, where $i := (r - j)/m$. Therefore, the distance from its centre to vertices is given by

$$d := (j3^{-2(i+1)} + (m-j)3^{-2i})^{\frac{1}{2}}/2, \tag{9}$$

which decreases with the increase in $i \in \mathbb{N}$. It is established in (Gablonsky, 2001, Thm. 4.2) that given any $i \in \mathbb{N}$, after $N \in \mathbb{N}$ iterations, DIRECT will leave no rectangle of side length greater than or equal to 3^{-i} in Ω, where N is as defined in (7) by Gablonsky (2001). When this is the case, in view of (9), no rectangle has a centre-to-vertex distance greater than $(m3^{-2i})^{\frac{1}{2}}/2$. Combining this with (8) and (6) implies at least one point on which Q evaluates to a value no greater than $y^* + \eta$ has been sampled by DIRECT. □

Remark 7 *The estimate (7) derived in (Gablonsky, 2001, Thm. 4.2) is based on the worst-case analysis which assumes that only one hyper-rectangle is subdivided per DIRECT iteration. In general, N in (7) is an upper*

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST

bound on the number of iterations needed before all rectangles' side lengths are less than 3^{-i}. The bound is tight, for example, when the objective function is constant.

3.2 Imprecise sampling

Robustness analysis of the DIRECT method is performed in this section, which has not appeared in the literature. This is exploited in establishing the main extremum seeking convergence result in the sequel. Frequently the output samples of the function Q are corrupted by some ν-bounded additive noise, i.e. given any $u \in \Omega$, the sample of Q evaluated at u can take any value in

$$Q(u) + [-\nu, \nu] := \{y \in \mathbb{R} \mid y = Q(u) + \delta;\ |\delta| \leq \nu\},$$

for some $\nu \geq 0$. As such, given any $u \in \Omega$ and $\tilde{Q} \in \mathbb{Q}(Q, \nu)$, we have that

$$|\tilde{Q}(u) - Q(u)| \leq \nu,$$

i.e. $\tilde{Q}$ represents Q perturbed by an additive noise whose magnitude is bounded above by ν.

Theorem 8 *Given a globally Lipschitz continuous function $Q : \Omega \subset \mathbb{R}^m \to \mathbb{R}$ with Lipschitz constant L and any $\tilde{Q} \in \mathbb{Q}(Q, \nu)$ for some $\nu \geq 0$, let*

$$y^* := \min_{u \in \Omega} Q(u).$$

Suppose for an $\eta > 0$, N and i are chosen such that (7) and (8) are satisfied. Denote by $\{\hat{y}_k\}_{k=1}^{\infty}$ the estimate sequence from DIRECT when applied to optimising $\tilde{Q}$. It holds that

$$|\hat{y}_N - y^*| \leq \eta + \nu.$$

PROOF. After N number of DIRECT iterations, let u be the centre point of any hyper-rectangle $H \subset \Omega$. We know by Lemma 6 that the minimum value Q can take on H is $Q(u) - \eta$. Furthermore, note that

$$\sup_{\tilde{Q} \in \mathbb{Q}(Q,\nu)} \left| \tilde{Q}(u) - (Q(u) - \eta) \right| \leq \nu + \eta.$$

The claimed result then follows from the inequality above, in light of the estimate update of DIRECT described in (4), which is set to be the minimum of $\tilde{Q}$ evaluated at the centres of all existing hyper-rectangles in Ω. □

We once again emphasise that N in Theorem 8 gives only an upper bound on the number of runs DIRECT needs before a sample of at most $\eta + \nu$ distance from y^* is taken. A criterion for when such a point is sampled is given by the next result.

Lemma 9 *Following Theorem 8, if for a $q \leq N$ the minimising input u_b which results in the current best estimate $\hat{y}_q = \tilde{Q}(u_b)$ is the centre point of a hyper-rectangle with centre-to-vertex distance d satisfying $Ld \leq \eta$, then it holds that*

$$|\hat{y}_q - y^*| \leq \eta + \nu,$$

and

$$|Q(u_b) - y^*| \leq \eta + 2\nu \tag{10}$$

PROOF. Let $Q_{max} : \Omega \to \mathbb{R}$ be defined by

$$Q_{max}(u) := Q(u) + \nu,$$

which is Lipschitz continuous with constant L. In effect, for any $u \in \Omega$, $Q_{max}(u)$ is the upper bound on the perturbed value of $Q(u)$ by ν-bounded noise. It follows by the same argument before (6) that $\hat{y}_q - Ld$ is a lower bound on $\min_{u \in \Omega} Q(u) = y^* + \nu$, i.e. $\hat{y}_q - Ld \leq y^* + \nu$. This implies that

$$|\hat{y}_q - y^*| \leq \eta + Ld \leq \eta + \nu.$$

That (10) holds then follows from $|\hat{y}_q - Q(u_b)| \leq \nu$ by the definition of $\tilde{Q}$. □

Remark 10 *It can be seen from the proof above that the lemma would still hold with the global Lipschitz constant L replaced with a local Lipschitz bound for a neighbourhood encompassing the hyper-rectangle in which the current best estimate $\hat{y}_q$ lies.*

Remark 11 *In the case where the Lipschitz constant L of the objective function Q is known, the DIRECT Algorithm 1 for a static map would normally be modified for efficiency by restricting the search for $\tilde{L}$ in (3) to those which satisfy $\tilde{L} \leq L$; this is not necessary for convergence. In the face of imprecise sampling of Q, i.e. sampling from a member $\tilde{Q}$ of $\mathbb{Q}(Q, \nu)$, this modification should* not *be enforced since there is no guarantee that $\tilde{Q}$ would be Lipschitz continuous with the same Lipschitz bound as that of Q. Otherwise, it may result in certain potentially optimal hyper-rectangles not being selected for subdivision and thus possible failure in locating a global extremum.*

4 Extremum seeking via the DIRECT method

4.1 The dynamical plant

This section demonstrates the application of the DIRECT method to seeking a global extremum of the

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST

steady-state behaviour of a MISO dynamical system with a compact input set through the use of periodic sampled-data components. In particular, consider the following definition of dynamical systems, which is largely based on that in Teel and Popović (2001).

In the definition below, $\mathcal{X}$ denotes a Banach space equipped with the norm $\|\cdot\|$.

Definition 12 *Let the state of a time-invariant dynamical system be represented by $x : \mathbb{R}_{\geq 0} \to \mathcal{X}$. The input to and output from the system are denoted, respectively, by $u : \mathbb{R}_{\geq 0} \to \Omega \subset \mathbb{R}^m$ and $y : \mathbb{R}_{\geq 0} \to \mathbb{R}$. Given any $u \in \Omega$ and $x_0 \in \mathcal{X}$, let $x(\cdot, x_0, u)$ denote the state of the dynamical system starting at x_0 with input u. The system is assumed to satisfy the following properties:*

(i) *There exists a function $\mathcal{A}$ mapping from Ω to subsets of $\mathcal{X}$ such that for each constant $u \in \Omega$, $\mathcal{A}(u)$ is a nonempty closed set and a global attractor (Ruelle, 1989). Furthermore,*

$$\sup_{u\in\Omega} \sup_{x\in\mathcal{A}(u)} \|x\| < \infty.$$

(ii) *There exists a locally Lipschitz function $h : \mathcal{X} \to \mathbb{R}$ such that the system output*

$$y(t) = h(x(t, x_0, u)) \quad \forall t \geq 0$$

for any constant input $u \in \Omega$ and $x_0 \in \mathcal{X}$. Moreover, $h(x_a) = h(x_b)$ for every $x_a, x_b \in \mathcal{A}(u)$. Since $\mathcal{A}(u)$ is a global attractor and h is locally Lipschitz, for any $u \in \Omega$ and $x_0 \in \mathcal{X}$,

$$\begin{aligned} Q(u) &:= \lim_{t\to\infty} h(x(t, x_0, u)) \\ &= h\left(\lim_{t\to\infty} x(t, x_0, u)\right) \\ &= h(x_l) \quad \textit{for any } x_l \in \mathcal{A}(u) \end{aligned}$$

is a well-defined readout map that is Lipschitz on Ω as per Assumption 1.

(iii) *For each triplet $(\epsilon_1, \epsilon_2, \Delta)$ of strictly positive real numbers, there exists a waiting time $T > 0$ such that if $\|x_0\|_{\mathcal{A}(u)} \leq \Delta$,*

$$\|x(t, x_0, u)\|_{\mathcal{A}(u)} \leq \epsilon_1 \|x_0\|_{\mathcal{A}(u)} + \epsilon_2 \quad \forall t \geq T, u \in \Omega.$$

Note that in the above definition, $u(t)$ is assumed to take values in a compact set $\Omega \subset \mathbb{R}^m$ for all $t \geq 0$. This arises most naturally from actuator constraints on the inputs in the form of saturation nonlinearity (Khalil, 2002).

Let $\{u_k\}_{k=0}^{\infty}$ be a sequence of vectors in Ω and define the zero-order hold (ZOH) operation

$$u(t) := u_k \quad \text{for all } t \in [kT, (k+1)T) \tag{11}$$

and $k = 0, 1, 2, \ldots$, where $T > 0$ denotes the period or waiting time. Let the state and output of a dynamical system in Definition 12 with respect to the input u be respectively x and y and define the periodic sampling operation $x_k := x(kT)$;

$$y_k := y(kT) \quad \text{for all } k = 1, 2, \ldots. \tag{12}$$

The following lemma is a generalisation of (Nešić et al., 2012, Prop. 1) to a broader class of dynamical systems modelled by Definition 12, including those of infinite dimension. The proof exploits the ideas in that of (Nešić et al., 2012, Prop. 1).

Lemma 13 *Given any dynamical system described in Definition 12, $\Delta > 0$, and $\nu > 0$, there exists a $T > 0$ such that for any $\{u_k\}_{k=0}^{\infty} \subset \Omega$ and $\|x_0\|_{\mathcal{A}(u_0)} \leq \Delta$,*

$$|y_k - Q(u_{k-1})| \leq \nu \quad \textit{for all } k = 1, 2, \ldots,$$

where y_k is as in (12) with y being the output of the system for the input u given by (11). Furthermore, given any $\hat{k} \in \mathbb{N}$ and $u_b \in \Omega$, suppose $u(t) := u_b \, \forall t \geq \hat{k}T$, then for any $\|x_0\|_{\mathcal{A}(u_0)} \leq \Delta$,

$$|y(t) - Q(u_b)| \leq \nu \quad \textit{for all } t \geq (\hat{k} + 1)T,$$

where $T > 0$ is the same waiting time as above.

PROOF. Let Δ and ν be given as in the lemma statement and L_h be the Lipschitz constant of h on the compact ball

$$\{x \in \mathcal{X} \mid \|x\| \leq \mathcal{A}_{max} + 1\},$$

where $\mathcal{A}_{max} := \sup_{u\in\Omega} \sup_{x\in\mathcal{A}(u)} \|x\|$, which is finite by Definition 12. Choose $\epsilon_1 > 0$ and $\epsilon_2 > 0$ so that

$$\epsilon_1(\Delta + 2\mathcal{A}_{max} + 1) + \epsilon_2 \leq \min\left\{\frac{\nu}{L_h}, 1\right\}.$$

By Property (iii) of Definition 12, it follows that there exists a $T > 0$ such that for any $\|x_0\|_{\mathcal{A}(u_0)} \leq \Delta + 2\mathcal{A}_{max} + 1$,

$$\begin{aligned} \|x_1\|_{\mathcal{A}(u_0)} &= \|x(T, x_0, u_0)\|_{\mathcal{A}(u_0)} \\ &\leq \epsilon_1 \|x_0\|_{\mathcal{A}(u_0)} + \epsilon_2 \\ &\leq \epsilon_1 \Delta + \epsilon_2 \\ &\leq 1, \end{aligned}$$

whereby $\|x_1\| \leq \mathcal{A}_{max} + 1$. We show below that $\|x_k\| \leq \mathcal{A}_{max} + 1$ for all $k = 1, 2, \ldots$ following an inductive argument. Suppose this is true for a $k \in \mathbb{N}$, which implies $\|x_k\|_{\mathcal{A}(u_k)} \leq 2\mathcal{A}_{max} + 1$. Then, by time-invariance of the

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST

dynamical system,

$$\begin{aligned}\|x_{k+1}\|_{\mathcal{A}(u_k)} &= \|x(T,x_k,u_k)\|_{\mathcal{A}(u_k)}\\ &\le \epsilon_1\|x_k\|_{\mathcal{A}(u_k)}+\epsilon_2\\ &\le \epsilon_1(2\mathcal{A}_{max}+1)+\epsilon_2\\ &\le 1,\end{aligned}$$

whereby $\|x_{k+1}\| \le \mathcal{A}_{max}+1$. Consequently,

$$\begin{aligned}|y_k - Q(u_{k-1})| &= \inf_{x_l\in\mathcal{A}(u_{k-1})}|h(x_k)-h(x_l)|\\ &\le L_h \inf_{x_l\in\mathcal{A}(u_{k-1})}\|x_k - x_l\|\\ &= L_h\|x_k\|_{\mathcal{A}(u_{k-1})}\\ &\le L_h(\epsilon_1\|x_{k-1}\|_{\mathcal{A}(u_{k-1})}+\epsilon_2)\\ &\le L_h(\epsilon_1(2\mathcal{A}_{max}+1)+\epsilon_2)\\ &\le \nu,\end{aligned}$$

where the first equality follows from Property (ii) of Definition 12 and L_h is as defined at the beginning of the proof. The last part of the lemma can be shown using the same argument as above. □

Remark 14 *The proof of Lemma 13 is based on the fact that the global asymptotic stability of dynamical systems stated in Definition 12 guarantees that after a sufficiently long waiting period T, the dynamics of the time-invariant system (13), initialised at any point in time with respect to any initial condition, will vanish to within a small perturbation on its steady-state input-output behaviour. Moreover, the magnitude of the perturbation reduces with the elongation of the waiting period T.*

A common class of systems (Ariyur and Krstić, 2003; Nešić et al., 2012; Tan et al., 2006) which satisfy Definition 12 is given by the following:

$$\begin{aligned}\dot{x} &= f(x,u) \quad x(0)=x_0;\\ y &= h(x),\end{aligned} \tag{13}$$

where $f:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ and $h:\mathbb{R}^n\to\mathbb{R}$ are locally Lipschitz functions in each argument.

Assumption 15 *There exists a locally Lipschitz function $\ell:\Omega\to\mathbb{R}^n$ such that*

$$f(\ell(u),u)=0 \quad \forall u\in\Omega.$$

Furthermore, $x=\ell(u)$ is globally asymptotically stable uniformly in $u\in\Omega$ (Khalil, 2002).

Definition 16 *Let*

$$Q(\cdot) := h\circ\ell(\cdot):\Omega\to\mathbb{R}$$

be the steady-state input-output (i.e. readout) map of system (13). Note that Q is globally Lipschitz on Ω because h is locally Lipschitz on $\mathbb{R}^n$ and ℓ is globally Lipschitz on Ω.

One may easily verify that the finite-dimensional nonlinear system (13) satisfies Definition 12, for which Lemma 13 holds. A less general version of the lemma for (13) can be found in (Nešić et al., 2012, Prop. 1).

4.2 Extremum seeking

Extremum seeking is about locating an extremal point of the steady-state input-output map Q of a dynamical system defined in Definition 12 (Ariyur and Krstić, 2003). Consider the modified-DIRECT-based extremum seeking scheme illustrated in Figure 1. The feedback system is interconnected through an ideal sampler of period $T>0$ (cf. (12)) and a synchronised zero-order hold (ZOH) (cf. (11)) device. There, the modified DIRECT algorithm is proposed in the following.

Algorithm 2 *The modified DIRECT for the extremum seeking setup in Figure 1 is as described in Algorithm 1, but with an additional input $\eta>0$ (the error margin) and an amendment to its termination criterion. In particular, Step* (vi) *of Algorithm 1 is revised to be the following.*

(vi) *Let d be the centre-to-vertex distance of the hyper-rectangle within which u_b satisfying $\hat{y}_q = Q(u_b)$ lies (cf. (4)). If $Ld\le\eta$, set the subsequent sample points to be the input which results in $\hat{y}_q$:*

$$u_{k+j} := u_b \quad \textit{for all } j=1,2,\ldots.$$

Otherwise, loop from (iii).

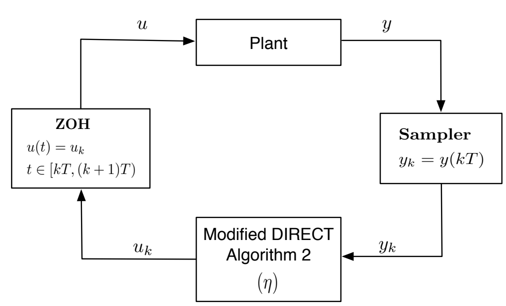

Fig. 1. Extremum seeking based on modified DIRECT

The main extremum seeking convergence result of the paper is given by the next theorem.

Theorem 17 *The closed-loop system depicted in Figure 1, consisting of a dynamical plant satisfying Definition 12, periodic sampler (12), zero-order hold (11), and modified DIRECT Algorithm 2 has the following convergence property: Given any $\Delta>0$ and $\mu>0$, let the parameter η of Algorithm 2 be any positive number less than*

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST

μ, then there exist a sampling/waiting period $T > 0$ and convergence time $\tilde{T} > 0$ such that for any $\|x_0\|_{\mathcal{A}(u_0)} \leq \Delta$,

$$|y(t) - y^*| \leq \mu \quad \text{for all } t \geq \tilde{T},$$

where $y^* := \min_{u \in \Omega} Q(u)$ and $Q : \Omega \subset \mathbb{R}^m \to \mathbb{R}$ denotes the steady-state input-output map with Lipschitz constant L as in Definition 12.

PROOF. Let $\nu := (\mu - \eta)/3$. Application of Lemma 13 to the plant with respect to ν yields a sampling/waiting period $T > 0$ such that for any $\|x_0\|_{\mathcal{A}(u_0)} \leq \Delta$,

$$|y_k - Q(u_{k-1})| \leq \nu \text{ for all } k = 1, 2, \ldots.$$

Correspondingly, there exists a $\tilde{Q} \in \mathbb{Q}(Q, \nu)$ for which the samples y_k are the evaluations of $\tilde{Q}$ on u_k. Theorem 8 then gives an upper bound $N \in \mathbb{N}$ such that there exists a $q \leq N$ for which the input u_b resulting in the current best estimate $\hat{y}_q$ of DIRECT Algorithm 2 lies in a hyper-rectangle of centre-to-vertex distance d that satisfies $Ld \leq \eta$, whereby invocation of Lemma 9 yields

$$|\hat{y}_q - y^*| \leq \nu + \eta$$

and

$$|Q(u_b) - y^*| \leq 2\nu + \eta. \tag{14}$$

Denote by $t_0 \geq 0$ the time at which the DIRECT iteration within which u_b is sampled ends. Note that by Step (vi) of DIRECT Algorithm 2,

$$u(t) := u_b \quad \text{for all } t \geq t_0.$$

Let $\tilde{T} := t_0 + T$. By the second part of Lemma 13, we have that

$$|y(t) - Q(u_b)| \leq \nu \quad \text{for all } t \geq \tilde{T}.$$

Together with (14), this implies that

$$|y(t) - y^*| \leq 3\nu + \eta = \mu \quad \text{for all } t \geq \tilde{T}. \qquad \square$$

Remark 18 *The convergence result in Theorem 17 is semi-global since Δ can be made arbitrarily large. It is also practical since μ can be arbitrarily small.*

Remark 19 *By Remark 10, the Lipschitz bound L used in Algorithm 2 can in general be less than that for Q on its whole domain Ω, while Theorem 17 still holds. It can be seen from Algorithm 2 that a large L leads to a higher number of DIRECT iterations before termination, and therefore slower speed of convergence.*

Remark 20 *Theorem 17 provides a convergence proof for an extremum seeking scheme for a general dynamical plant based on the DIRECT method (cf. Algorithm 2). The control scheme takes a finite amount of time to locate a reference input to the system which drives the steady-state behaviour into a neighbourhood of a global extremum. The region and accuracy of convergence can be improved at the expense of convergence rate, which corresponds to an increase in the waiting period T and/or number of iterations of DIRECT.*

Remark 21 *In the case where the dynamical system is slowly time-varying, the extremum seeking scheme depicted in Figure 1 can be 'restarted' after a period of idle time (during which Algorithm 2 persistently outputs the previously determined minimising command) to recalibrate the system and locate a (possibly different) global extremum of the steady-state. A sampled-data controller which cycles through several basic 'modes' of operation in a similar fashion described may be found, for e.g., in Nešić and Sontag (1998), where input-to-state stability in the presence of signed measurement disturbance is investigated.*

5 Extension to multiple units

It is common in engineering practice that the steady-state behaviour of a number of identical units needs to be optimised, for e.g., a solar cell array, production lines etc. (Woodward et al., 2009). When this is the case, exploitation of parallel computing to accelerate the search for an optimum becomes possible (Strongin and Sergeyev, 2000). In particular, each of these units may be fed with a different input simultaneously and the corresponding outputs collected/sampled after the waiting time, which amounts to performing several function evaluations of the (perturbed) steady-state behaviour within the same sampling period. The objective of this section is to generalise the sampled-data based global extremum seeking control scheme in Theorem 17 to the case involving multiple units.

The DIRECT Algorithm 2 is particularly well-suited for efficient parallel computing due to the following reason. During every iteration, there may be, depending on the objective function being optimised, a decently large number of potentially optimal hyper-rectangles (cf. Algorithm 1). In general, this may increase with the number of existing rectangles created by DIRECT in the search space $\Omega \subset \mathbb{R}^m$. Furthermore, several points that are aligned with the canonical basis vectors of $\mathbb{R}^m$ and of appropriate distance from the centre of each potentially optimal rectangle need to be sampled, if this was not done previously (cf. Algorithm 1(iv)). As such, the availability of multiple units can be utilised to perform the multi-sampling task. Recall from Figure 1 that the single-unit based extremum seeking scheme takes T seconds, i.e., the waiting time, to collect every sample se-

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST

quentially. Parallel computing thus allows several samples to be collected within a period of T seconds, as is detailed in the following figure.

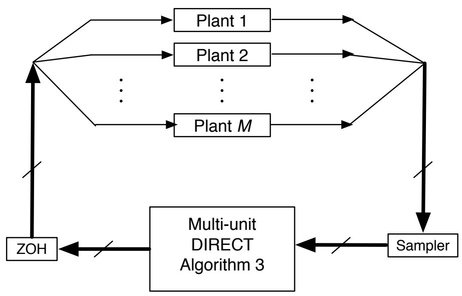

Fig. 2. Extremum seeking based on modified DIRECT

All dynamical plants in Figure 2 are assumed to satisfy Definition 12.

Remark 22 *It is apparent from the condition* (iii) *of the definition that the dynamics of each of the plant needs not be identical. They only need to decay at a sufficiently fast rate as determined by the waiting time T. We do assume that all plants share the same readout map, as defined in* (ii) *of Definition 12. Robustness of the proposed extremum seeking control framework to the converse would be an important issue worth investigating in future research.*

Figure 2 can be regarded as a master-slave paradigm, in which the multi-unit DIRECT algorithm is the master device which distributes the tasks to the plants, i.e. the slave processors, by appropriately selecting their inputs; see Watson and Baker (2001) more for details. In what follows, let M denote the total number of units/plants.

Algorithm 3 *The multi-unit DIRECT algorithm differs to Algorithm 2 only by the number of trial inputs to the multi-unit plants simultaneously. Let p be the number of samples DIRECT needs to collect during some iteration. If $p \leq M$, then all the units can be probed with appropriate inputs for this purpose. Otherwise, if $p > M$, by utilising the most number of plants available for data collection, it takes $r := \lceil p/M \rceil$ cycles/steps to sequentially collect the required samples, where $\lceil \cdot \rceil$ denotes the ceiling function. This amounts to rT seconds in the setup of Figure 2 per DIRECT iteration, where T denotes the sampling period.*

Using the same notation as the previous section, we have the following result.

Theorem 23 *Given any $\Delta > 0$ and $\mu > 0$, let the parameter $\eta > 0$ for Algorithm 3 be a number less than μ. There exist a sampling/waiting period $T > 0$ for the multi-unit extremum seeking scheme in Figure 2 and a convergence time $\tilde{T}_m > 0$ such that for any $\|x_0\|_{\mathcal{A}(u_0)} \leq \Delta$,*

$$|y(t) - y^*| \leq \mu \quad \textit{for all } t \geq \tilde{T}_m,$$

where y denotes the output of any of the unit/system. Furthermore, suppose $\tilde{T}$ is a convergence time with respect to which the single-unit extremum seeking based Theorem 17 holds, then by employing multiple units it holds in general that $\tilde{T}_m \leq \tilde{T}$, i.e. the multi-unit extremum seeking scheme converges to a global minimum no slower than the single-unit scheme.

SKETCH OF PROOF. By hypothesis and Lemma 13 a waiting time T_i can be selected so that the difference at sampling between the output of the i^{th} dynamical unit and its corresponding steady-state value is no larger than $\nu := (\mu - \eta)/3$ for every $i = 1, \ldots, M$. Set

$$T := \max_{i=1,\ldots,M} T_i.$$

The remaining proof for convergence is then the same as that for Theorem 17.

The last part of theorem is straightforward since probing multiple units concurrently as opposed to a single one results in the generation of more sampled values after each waiting period. □

Remark 24 *In the case of serial optimisation, every input trial point determined by an algorithm for the objective function is based on all the information obtained hitherto. In contrast, when a parallel computing scheme is deployed, each of the $p > 1$ trial points within an iteration is chosen in the absence of information regarding the function evaluation results of the others, and hence may itself be redundant, thereby reducing the efficiency of the algorithm. This problem of* redundancy *of samples in parallelism, as described in (Strongin and Sergeyev, 2000, Chapter 5), is in fact* not *a concern for DIRECT. This is because the standard DIRECT algorithm is set up so that within every iteration, several samples which correspond to neighbouring points of potentially optimal hyperrectangles may need to be taken. These potentially optimal rectangles are determined by the partitioning which took place in the previous iteration, and are therefore not dependent on each other.*

6 Numerical simulations

The simulation results in this section are generated with the aid of the MATLAB source code from Finkel (2004).

6.1 Single unit

We consider the following two-dimensional dynamical system with two inputs which is of the differential form

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST

in (13):

$$\begin{aligned} \dot{x}_1 &= -2x_1 + u_1, \qquad && x_1(0) := 2; \\ \dot{x}_2 &= x_1 - x_2^3 + u_2, && x_2(0) := -1; \\ y &= h(x), \end{aligned}$$

where for all $t \geq 0$,

$$u(t) \in \Omega := \left\{ u \in \mathbb{R}^2 : -5 \leq u_1 \leq 10,\ 0 \leq u_2 \leq 15 \right\}$$

and

$$h(x_1, x_2) := \left(x_2^3 - 2 - \frac{5.1}{\pi^2} x_1^2 + \left(\frac{10}{\pi} - \frac{1}{2} \right) x_1 - 6 \right)^2 + 10 \left(1 - \frac{1}{8\pi} \right) \cos 2x_1 + 10.$$

It follows that for any $u \in \Omega$, $x_1 = 0.5u_1$ and $x_2 = (0.5u_1 + u_2)^{\frac{1}{3}}$ is a globally asymptotically stable equilibrium. Thus, the steady-state map (cf. Definition 16) of the above dynamical system is given by

$$Q(u) = \left(u_2 - 2 - \frac{5.1}{4\pi^2} u_1^2 + \frac{5}{\pi} u_1 - 6 \right)^2 + 10 \left(1 - \frac{1}{8\pi} \right) \cos u_1 + 10,$$

which is the so-called Branin test function for bound-constrained optimisation (Branin, 1972; Jones et al., 1993); see Figure 3 for plots of the function. The global minimum value of Q is 0.3978 and the Lipschitz bound is in the order of 100.

With $L := 5$ and $\eta := 0.01$, Figure 4 shows the current estimate of the minimum of Q found by DIRECT over time within the sampled-data based extremum seeking framework of Figure 1 for two different sampling/waiting periods, namely $T = 0.1$s and $T = 0.5$s. Figure 5 shows the corresponding outputs of the system. Recall from Lemma 13 that the use of a smaller waiting time results in lesser estimation accuracy. This fact is summarised in Table 1, in which for the case $T = 0.5$s the final estimate is fairly close to the global minimum value at the expense of longer convergence duration. In reference to Remarks 10 and 19, the latter example also demonstrates that an exact global Lipschitz constant for the objective function is not always needed to obtain an accurate extremum estimate. Note that the DIRECT iteration number does not necessarily decrease with the increase of the waiting time since modifying the waiting time changes the directions in which the dynamics perturb the steady-state input-output map at the instants sampling takes place. The additional iterations and samples for the $T = 0.5$s case were needed in order to locate a global minimum.

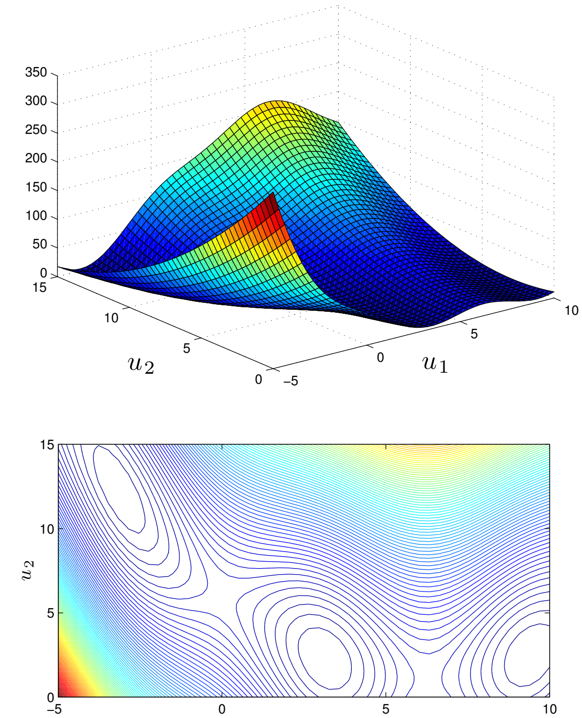

Fig. 3. The Branin function. (Top: surface plot; bottom: contour plot.)

Table 1
Performance characteristics

T	Est. min	Duration	Iter. no.	Sample no.
0.1s	1.193	8.1s	9	81
0.5s	0.401	77.5s	14	155

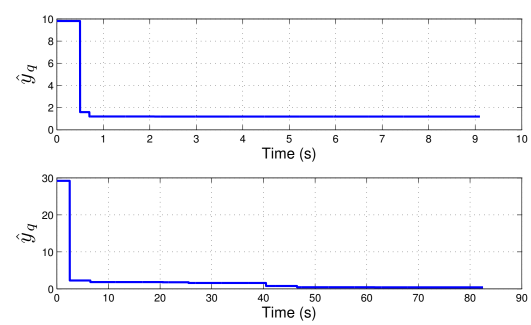

Fig. 4. The estimates by DIRECT over time for waiting periods $T = 0.1$s (top) and $T = 0.5$s (bottom).

Finally, by setting $T := 0.5$s and $L := 2$ and re-running extremum seeking, DIRECT terminates at its 8^{th} iteration or 32.5s and yields an inaccurate estimate of the function's minimum value of 1.563. This complies with Algorithm 2 since the Lipschitz bound L determines when the modified DIRECT halts and the use of a bound of smaller value gives rise to earlier termination of DI-

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST

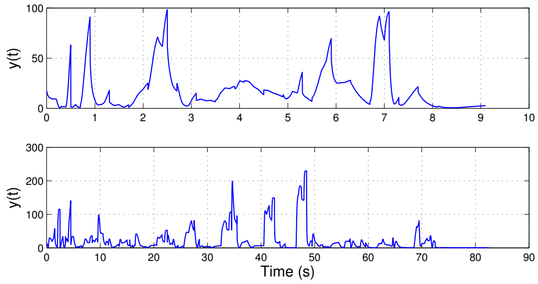

Fig. 5. System's output for waiting periods $T = 0.1$s (top) and $T = 0.5$s (bottom).

RECT and larger convergence error; see Figures 6 and 7.

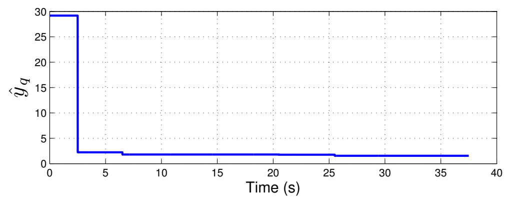

Fig. 6. The estimate by DIRECT for $L = 2$ ($T = 0.5$s).

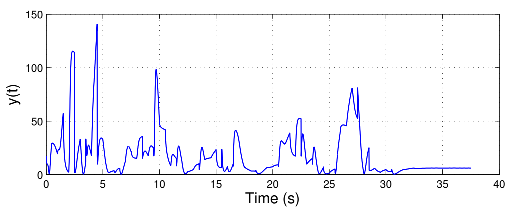

Fig. 7. System's output for $L = 2$ ($T = 0.5$s).

6.2 *Multiple units*

The number of samples taken during the 14 iterations of DIRECT for the case $T = 0.5$s and $L = 5$ considered in the previous subsection are, respectively, 5, 2, 6, 6, 8, 14, 10, 14, 16, 12, 18, 12, 14, and 18. These total to 155, as shown in the last entry of Table 1.

In view of the multi-unit extremum seeking control framework of Figure 2 and Algorithm 3, if the number of available plants is $M \geq 20$, then it takes only $14 \times 0.5\text{s} = 7\text{s}$ to locate the value 0.401, by contrast with 77.5s in the second row of Table 1. On the other hand, if $M = 18$, then the 14 iterations respectively take the following multiples of the waiting period to complete: 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3, 2, 2, 3. As such, $25 \times 0.5\text{s} = 12.5\text{s}$ is needed to conclude all the iterations before the value 0.401 is found. These are illustrated in Figure 8.

In general, employment of multiple units for parallel computing in extremum seeking can significantly accelerate the search, as justified by the analysis above. This owes to the fact emphasised earlier that there could be more than 1 potentially optimal hyper-rectangle during

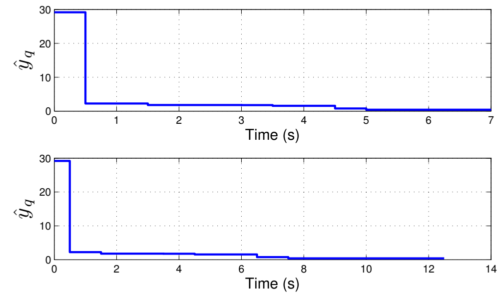

Fig. 8. The estimates by DIRECT over time for unit number $M = 18$ (top) and $M = 8$ (bottom).

a DIRECT iteration and sampling is done in all dimensions of the search domain .

7 Concluding remarks

A sampled-data extremum seeking control of dynamical systems is proposed based on a modified DIRECT global optimisation method. We establish semi-global practical convergence of the scheme with respect to two design parameters, the waiting/sampling period T and DIRECT-termination accuracy parameter η. A multi-unit scheme is also proposed to expedite the convergence speed to a global extremum. Future research directions involve adapting other sampling-based optimisation algorithms in Strongin and Sergeyev (2000), for instance, within the context of extremum seeking control and developing a unifying framework for a wide class of algorithms in the spirit of Nešić et al. (2010); Teel and Popović (2001).

References

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

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization.* Cambridge University Press.

Branin, F. H. (1972). Widely convergent method for finding multiple solutions of simultaneous nonlinear equations. *IBM Journal of Research and Development*, 16(5):504–522.

Esmaeil-Zadeh-Azar, F. (2010). *An Efficient Global Optimization Method Based on Multi-Unit Extremum Seeking.* PhD thesis, University of Montréal.

Finkel, D. E. (2004). DIRECT research and codes. `http://www4.ncsu.edu/~ctk/Finkel_Direct/`.

Finkel, D. E. (2005). *Global Optimization with the DIRECT Algorithm.* PhD thesis, North Carolina State University.

Gablonsky, J. M. (2001). *Modificaions of the DIRECT algorithm.* PhD thesis, North Carolina State University.

Gablonsky, J. M. and Kelley, C. T. (2001). A locally-biased form of the DIRECT algorithm. *Journal of Global Optimization*, 21(1):27–37.

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST

CONFIDENTIAL. Limited circulation. For review only

Graham, R. L. (1972). An efficient algorithm for determining the convex hull of a finite planar set. *Information Processing Letters*, 1:132–133.

Jones, D. R., Pertunnen, C. D., and Stuckman, B. E. (1993). Lipschitzian optimization without the Lipschitz constant. *Journal of Optimization Theory and Applications*, 79(1):157–181.

Khalil, H. K. (2002). *Nonlinear Systems*. Prentice Hall, 3rd edition.

Nešić, D., Nguyen, T., Tan, Y., and Manzie, C. (2012). A non-gradient approach to global extremum seeking: an adaptation of the Shubert algorithm. *Automatica*. In press.

Nešić, D. and Sontag, E. D. (1998). Input-to-state stabilization of linear systems with positive outputs. *Systems & Control Letters*, 35:245–255.

Nešić, D., Tan, Y., Moase, W. H., and Manzie, C. (2010). A unifying approach to extremum seeking: adaptive schmes based on estimation of derivatives. In *Proceedings of IEEE Conference on Decision and Control*, pages 4625–4630.

Pintér, J. D. (1996). *Global optimization in action: Continuous and Lipschitz optimization: Algorithms, implementations and applications*. Kluwer Academic Publishers.

Piyavskij, S. A. (1972). An algorithm for finding the absolute minimum of a functions. *USSR Computational Mathematics and Mathematical Physics*, 12:57–67.

Rudin, W. (1976). *Principles of Mathematical Analysis*. McGraw-Hill, 3rd edition.

Ruelle, D. (1989). *Elements of Differentiable Dynamics and Bifurcation Theory*. Academic Press.

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

Strongin, R. G. and Sergeyev, Y. D. (2000). *Global Optimization with Non-Convex Constraints*. Nonconvex optimization and its applications. Kluwer Academic Publishers.

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

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

Watson, L. T. and Baker, C. A. (2001). A fully distributed parallel global search algorithm. *International Journal of Computer Aided Engineering*, 18:155–169.

Woodward, L., Perrier, M., and Srinivasan, B. (2009). Improved performance in the multi-unit optimization method with non-identical units. *Journal of Process Control*, 19:205–215.

Zhang, C. and Ordóñez, R. (2011). *Extremum-Seeking Control and Applications: A Numerical Optimization-Based Approach*. Advances in Industrial Control. Springer.

Preprint submitted to Automatica
Received March 5, 2013 05:01:39 PST