Unified frameworks for sampled-data extremum seeking control: global optimisation and multi-unit systems ⋆

Sei Zhen Khong [a,1], Dragan Nešić [a], Ying Tan [a], Chris Manzie [b]

[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

Two frameworks are proposed for extremum seeking of general nonlinear plants based on a sampled-data control law, within which a broad class of nonlinear programming methods is accommodated. It is established that under some generic assumptions, semi-global practical convergence to a global extremum can be achieved. In the case where the extremum seeking algorithm satisfies a stronger asymptotic stability property, the converging sequence is also shown to be stable using a trajectory-based proof, as opposed to a Lyapunov-function-type approach. The former is more straightforward and insightful. This allows for more general optimisation algorithms than considered in existing literature, such as those which do not admit a state-update realisation and/or Lyapunov functions. Lying at the heart of the analysis throughout is robustness of the optimisation algorithms to additive perturbations of the objective function. Multi-unit extremum seeking is also investigated with the objective of accelerating the speed of convergence.

Key words: Extremum seeking, sampled-data control, nonconvex global optimisation, robustness, multi-unit systems

1 Introduction

Modelling of complicated nonlinear dynamical systems is often a challenging task. Many of such systems manifest an extremal operating condition in its steady-state input-output behaviour. In the absence of the knowledge of a model of the plant and its steady-state input-output map, *extremum seeking* is a real-time optimisation method that drives the system into a vicinity of this extremum (Ariyur and Krstić, 2003; Zhang and Ordóñez, 2011). Extremum seeking has found use in a broad array of applications, ranging from biochemical reactors (Wang et al., 1999a; Guay et al., 2003) and gas-turbine combustors (Moase et al., 2010) to axial compressors (Wang et al., 1999b) and optical fibre amplifiers (Dower et al., 2007).

Teel and Popović made a significant contribution to the area of extremum seeking in Teel and Popović (2001); Popović (2004), where it is shown that under assumptions on the asymptotic stability of both the plant and discrete-time nonlinear programming method, extremum seeking can be achieved within a periodic sampled-data framework. This gives rise to powerful capacity to utilise a wide class of optimisation algorithms for the task of steady-state extremum seeking of a dynamical system. In particular, complexity of implementation and convergence speed of the algorithms may be taken into account in the control design stage. An important feature of the framework of Teel and Popović (2001) is the online approximation of the derivative of the objective function by applying constant probing inputs to the plant successively. A Lyapunov stability proof of the scheme based on interconnected systems' theory is examined in Kvaternik and Pavel (2011) using stronger conditions.

In contrast to Teel and Popović (2001); Popović (2004), *adaptive* control methods have traditionally been the core of many extremum seeking schemes (Krstić and Wang, 2000; Ariyur and Krstić, 2003, 2004; Tan et al., 2006; Nešić, 2009; Ghaffari et al., 2012). In the absence of a model for a dynamical plant, the gradient of its steady-state map is adaptively estimated in real time using appropriate dither signals (Tan et al., 2008). Cor-

⋆ This work was supported by the Australian Research Council (DP0985388). This paper was not presented at any IFAC meeting.

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

[1] Tel. +61 3 83443934. Fax +61 3 83447412.

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

respondingly, local (Krstić and Wang, 2000) and semi-global (Tan et al., 2006) *asymptotically stable* convergence to an extremum (characterised by a zero derivative) is demonstrated by exploiting the time-scale separation between the dynamics of the plant and estimator via the methods of averaging and singular perturbation. Inspired by Teel and Popović (2001); Popović (2004), a systematic derivative-based framework for adaptive extremum seeking control is prescribed in Nešić et al. (2010, 2012) using similar techniques, in which a large class of continuous-time optimisation algorithms may be applicable. Related work on plants with parametric uncertainties is considered in Nešić et al. (2013a). These parallel the discrete-time counterpart of Teel and Popović (2001), where Lyapunov-based analysis methods are employed to establish asymptotic convergence. Note that Nešić et al. (2010, 2012) deal only with nonlinear dynamical systems that are of a finite dimension to which averaging and singular perturbation analysis (Khalil, 2002; Tan et al., 2006) applies, whereas infinite-dimensional systems can be accommodated in Teel and Popović (2001) and this paper.

Notice that in spite of the generality of the frameworks proposed in Teel and Popović (2001); Nešić et al. (2010), they both rely upon *gradient*-based optimisation with a state-update model. A well-known problem with optimisation methods of this sort is their inability to locate a *global* extremum amongst local ones (Luenberger, 1969; Boyd and Vandenberghe, 2004), and hence the convexity of the objective function plays a crucial role in determining the success of global extremum seeking. In contrast, many discrete-time sampling-based methods capable of non-convex optimisation are available in the literature (Pintér, 1996; Strongin and Sergeyev, 2000). These algorithms commonly cannot be put into the state-update form stipulated in Teel and Popović (2001); Nešić et al. (2010), nor do they possess the stability properties required. Yet recently, a weaker type of convergence to a global optimum is shown to be accomplishable based on the sampled-data framework of Teel and Popović (2001) for two algorithms of this sort, namely the Piyavskii-Shubert (Nešić et al., 2013b) and DIRECT (Khong et al., 2013). This motivates the development of a more fundamental framework within which to encompass these global *nonconvex* and nonsmooth optimisation algorithms.

This paper develops a unifying framework for a class of sampled-data extremum seeking controllers based on the notion of *attractivity* as opposed to asymptotic stability. It is shown that this covers an even larger class of extremum seeking controllers than those considered in Teel and Popović (2001), including sampling-type global algorithms. Semi-global practical convergence is established for the class of extremum seekers which satisfy a robustness property applied to asymptotically stable, possibly distributed-parameter, nonlinear dynamical plants. Examples of extremum seeking algorithms satisfying the required robustness property are given, which include the aforementioned Piyavskii-Shubert and DIRECT. In particular, a third example, the Global Search Algorithm (Strongin and Sergeyev, 2000), is examined in detail; this has not been done in the literature.

When the extremum seeking controllers possess a stronger asymptotically stable property, the converging sequence is also shown to be asymptotically stable using the notion of multi-step consistency/robustness. The proof provided is trajectory-based. It serves as a straightforward alternative to the Lyapunov-based proof given in Teel and Popović (2001); Popović (2004), which exploits the closeness of solutions to a differential inclusion form over a single time step. In particular, it is demonstrated that by restricting the abstract trajectory-based convergence result of this paper to one where the extremum seeking algorithms take a differential inclusion form as in Teel and Popović (2001); Popović (2004), the underlying technical assumptions employed herein are no stronger than those in Teel and Popović (2001); Popović (2004). Not only are the trajectory-based assumptions stated here generally easier to verify (or speculated via simulations/experiments), the proof of the main result, having no recourse to Lyapunov-type arguments, is also simpler than that in Teel and Popović (2001); Popović (2004). It is more insightful in the authors' beliefs, giving rise to the ideas underpinning the attractivity-based general framework described above. To use the results in Teel and Popović (2001); Popović (2004), one would typically need to construct a Lyapunov function or resort to converse Lyapunov theorems (Khalil, 2002; Kellet and Teel, 2005), whereas the results in this paper apply directly once asymptotic stability of extremum seeking algorithms is established.

Efficiency of extremum seeking can be increased by exploiting parallelism in computations (Strongin and Sergeyev, 2000). This paper considers generalisations of the two aforementioned frameworks to cases where multiple plants of similar but *non-identical* dynamics are available for probing as a means of accelerating the speed of convergence. Methods which do not give rise to redundancy in the sampling points are discussed for both the attractivity and asymptotic stability extremum seeking frameworks.

The paper is organised as follows. The type of dynamical plants considered for extremum seeking is described in the next section. A generic attractivity robustness property of extremum seekers is stated in Section 3, which is exploited to establish the main convergence proof for a sampled-data extremum seeking control framework in Section 4. In Section 5, a trajectory-based proof for asymptotically stable convergence is given and related with the more general framework in Section 4. A discussion about the relation with the work by Teel and Popović (2001) is provided in Section 6. The develop-

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

ments are furnished with several intertwining examples which meet the assumptions used in this paper. Extremum seeking with multiple units/plants is considered in Section 7 and some simulation examples in Section 8.

2 Dynamical systems

The class of nonlinear, possibly infinite-dimensional, systems considered in this paper is introduced in this section. Specific examples of such systems are presented. The following notational definitions are important.

A function $\gamma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class-$\mathcal{K}$ (denoted $\gamma \in \mathcal{K}$) if it is continuous, strictly increasing, and $\gamma(0) = 0$. If γ is also unbounded, then $\gamma \in \mathcal{K}_\infty$. A continuous function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class-$\mathcal{KL}$ if for each fixed t, $\beta(\cdot, t) \in \mathcal{K}$ and for each fixed s, $\beta(s, \cdot)$ is decreasing to zero (Khalil, 2002). The Euclidean norm is denoted $\|\cdot\|_2$.

Let $\mathcal{X}$ be a Banach space whose norm is denoted $\|\cdot\|$. Given any subset $\mathcal{Y}$ of $\mathcal{X}$ and a point $x \in \mathcal{X}$, define the distance of x from $\mathcal{Y}$ as $\|x\|_\mathcal{Y} := \inf_{a \in \mathcal{Y}} \|x - a\|$. Also let

$$\mathcal{U}_\epsilon(\mathcal{Y}) := \{x \in \mathcal{X} \mid \|x\|_\mathcal{Y} < \epsilon\}.$$

Definition 1 *Let the state of a time-invariant dynamical system be represented by $x : \mathbb{R}_{\geq 0} \to \mathcal{X}$, where $\mathcal{X}$ is a Banach space with norm $\|\cdot\|$. The input to and output of the system are denoted, respectively, by $u : \mathbb{R}_{\geq 0} \to \Omega \subset \mathbb{R}^m$ and $y : \mathbb{R}_{\geq 0} \to \mathbb{R}$. The set Ω denotes the input space of interest, and is taken to be a compact*[2] *subset of $\mathbb{R}^m$ in this paper. Given any $u \in \Omega$ and $x_0 \in \mathcal{X}$, let $x(\cdot, x_0, u)$ be the state of the dynamical system starting at x_0 with input u.*

The following assumption is based to a large extent on (Teel and Popović, 2001, Assumption 1).

Assumption 2 *Given a system described in Definition 1, the following hold:*

(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):*
  - *(a) Given any $x_0 \in \mathcal{X}$ and $\epsilon > 0$, there exists a sufficiently large $t > 0$ such that $x(t, x_0, u) \in \mathcal{U}_\epsilon(\mathcal{A}(u))$;*
  - *(b) If $x(t_0, x_0, u) \in \mathcal{A}(u)$, then $x(t, x_0, u) \in \mathcal{A}(u)$ for all $t \geq t_0$;*
  - *(c) There exists no proper subset of $\mathcal{A}(u)$ having the first two properties above.*

  Furthermore,

$$\sup_{u \in \Omega} \sup_{x \in \mathcal{A}(u)} \|x\| < \infty. \tag{1}$$

(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 $x_0 \in \mathcal{X}$,

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

is a well-defined steady-state input-output map that is Lipschitz on Ω.

(iii) *Q takes its global minimum value in a nonempty, compact set $\mathcal{C} \subset \Omega$.*

(iv) *Given any $\Delta > 0$, there exists a class-$\mathcal{KL}$ function β such that*

$$\|x(t, x_0, u)\|_{\mathcal{A}(u)} \leq \beta(\|x_0\|_{\mathcal{A}(u)}, t)$$

for all $t \geq 0$, $u \in \Omega$, and $\|x_0\|_{\mathcal{A}(u)} \leq \Delta$.

Remark 3 *The thesis by Popović (2004) considers also a closed subset of $\mathcal{X}$ on which the system is not allowed to operate. Such a consideration can be straightforwardly accommodated in this paper by strengthening Assumption 2 according to (Popović, 2004, Assumption 2.4(4)).*

In the following, two common classes of systems which satisfy Definition 1 and Assumption 2 are listed.

2.1 Systems with equilibria

Consider the following dynamical system (Ariyur and Krstić, 2003; Tan et al., 2006):

$$\begin{aligned} \dot{x} &= f(x, u) \quad x(0) = x_0; \\ y &= h(x), \end{aligned} \tag{2}$$

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

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

[2] This is not a stringent assumption given the ubiquity of control input saturation constraints in physical systems (Khalil, 2002).

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

Furthermore, $x = \ell(u)$ is globally asymptotically stable uniformly in $u \in \Omega$ (Khalil, 2002), i.e. there exists a $\beta \in \mathcal{KL}$ such that for any $u \in \Omega$ and $x_0 \in \mathbb{R}^n$,

$$\begin{aligned}\|x(t, x_0, u)\|_{\ell(u)} &= \|x(t, x_0, u) - \ell(u)\|_2 \\ &\leq \beta(\|x_0 - \ell(u)\|_2, t) \quad \forall t \geq 0,\end{aligned}$$

where $x(\cdot, x_0, u)$ denotes the solution to (2) with respect to the initial condition x_0 and input u.

Definition 5 *Let*

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

be the steady-state input-output map of system (2).

Suppose Q achieves its global minimum value on a nonempty compact set $\mathcal{C} \subset \Omega$. It can be verified easily that the system (2) satisfies Assumption 2.

2.2 Systems with periodic attractors

The following class of systems is adopted from Haring et al. (2012). Stronger differentiability assumptions are made in Haring et al. (2012), as required to apply the proof technique to conclude stability of the continuous-time extremum seeking scheme therein.

Consider a nonlinear plant model:

$$\begin{aligned}\dot{x} &= f(x, u, w) \quad x(0) = x_0; \\ y_p &= g(x, w),\end{aligned} \tag{3}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}^n$ and $g : \mathbb{R}^n \times \mathbb{R}^l \to \mathbb{R}$ are locally Lipschitz functions in each argument and w is T_w-periodic in time, i.e. $w(t+T_w) = w(t)$ for all $t \geq 0$.

Assumption 6 *For each fixed $u \in \Omega$, there exists a unique, bounded, uniformly globally asymptotically stable T_w-periodic solution $\hat{x}_u$ (Khalil, 2002) to (3) given by*

$$\hat{x}_u(t) = M(u, w(t)) \quad t \geq 0,$$

where $M : \Omega \times \mathbb{R}^l \to \mathbb{R}^n$ is locally Lipschitz in each argument.

It follows from the assumption above that the steady-state output of the plant (3) for a fixed $u \in \Omega$, i.e. $y_p = g(M(u, w), w)$, is a periodic, continuous function with period T_w. Define the performance measures:

$$\begin{aligned}L_p(y_p) &:= \left(\frac{1}{T_w} \int_0^{T_w} |y_p(t)|^p \, dt \right)^{\frac{1}{p}}, \quad p \in [1, \infty); \\ L_\infty(y_p) &:= \max_{t \in [0, T_w)} |y_p(t)|\end{aligned} \tag{4}$$

and the cost function $C_i(y_p) := h \circ L_i(y_p), i \in [1, \infty]$, where $h : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is a locally Lipschitz user-designed function. The cascade connection of the plant (3) and C_i for $i \in [1, \infty]$ admits the following steady-state map for $u \in \Omega$:

$$\begin{aligned}Q_p(u) &:= h \left(\left(\frac{1}{T_w} \int_0^{T_w} |g(M(u, w(t)), w(t))|^p \, dt \right)^{\frac{1}{p}} \right); \\ Q_\infty(u) &:= h \left(\max_{t \in [0, T_w)} |g(M(u, w(t)), w(t))| \right),\end{aligned}$$

where $p \in [1, \infty)$. Suppose Q_i takes its global minimum value on a nonempty compact $\mathcal{C} \subset \Omega$. It can be seen that the cascade of the plant and C_i satisfies Assumption 2 for $i \in [1, \infty]$.

3 Extremum seeking controllers

In this section, a generic robustness property of extremum seeking algorithms to bounded additive perturbations of measurements of the objective function is stated. This will later prove useful in establishing the convergence of extremum seeking in the next section. Three algorithms which are known to possess this property are given at the end.

Consider the optimisation problem:

$$y^* := \min_{u \in \Omega} Q(u), \tag{5}$$

where $Q : \Omega \subset \mathbb{R}^m \to \mathbb{R}$ is a Lipschitz continuous function which takes its global minimum value on $\mathcal{C} \subset \Omega$, i.e. $Q(u) = y^*$ for all $u \in \mathcal{C}$. Let Σ be a discrete-time extremum seeking algorithm for (5). The output sequence Σ generates to probe Q is denoted $\{u_k\}_{k=0}^\infty$ and in the case of precise (i.e. noiseless) sampling, the collected measurements are $y_k = Q(u_{k-1})$, $k = 1, 2, \ldots$. Define also the sequence

$$\bar{y}_N := \min_{k=1,\ldots,N} y_k. \tag{6}$$

Let $\hat{\delta}$ be a non-negative real number. It follows from the above definition that the sequence $\{\bar{y}_k\}_{k=1}^\infty$ converges to the closed $\hat{\delta}$-neighbourhood ($\hat{\delta}$-ball) of y^* if for all $\epsilon > 0$, there exists infinitely many $N \in \mathbb{N}$ such that

$$u_N \in \{u \in \Omega \mid |Q(u) - y^*| \leq \hat{\delta} + \epsilon\}.$$

By the Lipschitz continuity of Q, corresponding to the $\hat{\delta}$ above, there exists a $\delta > 0$ such that the aforementioned condition holds if the sequence $\{u_k\}_{k=0}^\infty$ converges to the δ-neighbourhood of $\mathcal{C}$. That is, for all $\epsilon > 0$, there exists

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

an $N \in \mathbb{N}$ such that

$$u_k \in \mathcal{C} + (\delta + \epsilon)\bar{\mathcal{B}} \quad \forall k \geq N, \tag{7}$$

where $\bar{\mathcal{B}}$ denotes the closed unit ball in $\mathbb{R}^m$.

In the presence of bounded additive perturbations on the measurements as illustrated in Figure 1, i.e. $y_k = Q(u_{k-1}) + w_k$ with $|w_k| \leq \nu$ for some $\nu > 0$, the following assumption is important to establish convergence of the sampled-data extremum seeking scheme to be considered in Section 4.

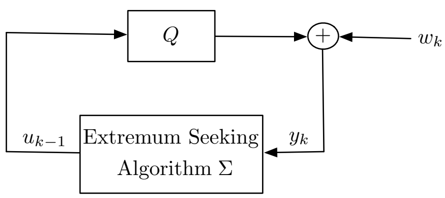

Fig. 1. Extremum seeking algorithm with noisy output measurement.

Assumption 7 *Let $\delta \geq 0$ be a small number which characterises the accuracy of convergence as in (7). The discrete-time extremum seeking algorithm Σ satisfies the following: Given any $\mu > \delta$, there exists a $\nu > 0$ such that if $|\hat{y}_k - Q(\hat{u}_{k-1})| \leq \nu$ for $k = 1, 2, \ldots$ and any sequence $\{\hat{u}_k\}_{k=0}^{\infty} \subset \Omega$, then for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ for which $u_k \in \mathcal{C} + (\mu + \epsilon)\bar{\mathcal{B}}$ for all $k \geq N$. In other words, the output of Σ, $\{u_k\}_{k=0}^{\infty}$ converges to a μ-neighbourhood of the set $\mathcal{C}$ of global minimisers of Q.*

In the following, three sampling-based algorithms that satisfy the assumption above are given. They do not require explicit estimation of derivatives of the steady-state map and are capable of locating a global optimum in the presence of local extrema.

3.1 The Piyavskii-Shubert method

The Piyavskii-Shubert method (Piyavskii, 1972; Shubert, 1972) is a global optimisation method which is well-suited for a compact one-dimensional input space, specifically a closed interval $\Omega := [a, b]$. Specifically, the algorithm is described by $u_0 := (b-a)/2$ and for $k = 0, 1, \ldots$,

$$u_{k+1} = \arg\min_{u \in [a,b]} \max_{j=1,\ldots,k} \{y_j - L|u - u_j|\}, \tag{8}$$

where L is the Lipschitz bound for the objective function $Q : \Omega \to \mathbb{R}$. Robustness analysis of the Piyavskii-Shubert method in Nešić et al. (2013b) shows that with $\delta = 0$ and given any $\mu > 0$, it satisfies Assumption 7 with $\nu = \mu/3$.

3.2 The DIRECT method

The DIRECT algorithm (Jones et al., 1993) addresses the curse-of-dimensionality drawback of Piyavskii-Shubert. Operating on a compact bound-constrained multi-dimensional domain of search

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

it is intelligently balanced between local and global search. The trial points DIRECT samples in the input space always form a dense subset, whereby the output sequence of DIRECT does not converge to a vicinity of a global extremum but only a subsequence of it does. This renders DIRECT unsuitable for the purpose of extremum seeking control. To circumvent this undesired property, Khong et al. (2013) proposes a DIRECT algorithm with a modified termination criterion where $\delta > 0$ in Assumption 7 is a design parameter. Given any $\mu > 0$, choose a positive $\delta < \mu$, then the modified DIRECT is shown in Khong et al. (2013) to satisfy Assumption 7 with $\nu = (\mu - \delta)/3$. The resulting sampled-data control scheme is of a type where the steady-state output of the plant is driven to a neighbourhood of its global extremum within finite time and the corresponding input is maintained from then onwards.

3.3 The Global Search Algorithm

The core Global Search Algorithm (GSA) from (Strongin and Sergeyev, 2000, Section 3.1) is considered and its robustness to additive perturbations on measurements analysed; the main result in this subsection is new. By contrast to the geometric procedures such as Shubert and DIRECT, the GSA is an information approach which bases its ideas on approximate stochastic estimators (Strongin and Sergeyev, 2000, Section 2.2). (Strongin and Sergeyev, 2000, Section 4.5) shows that the GSA possesses a geometric local tuning interpretation, i.e. adaptive estimation of local Lipschitz constants of the objective function. It thus sits between Shubert and DIRECT methods, the former of which assumes knowledge of the global Lipschitz constant while the latter examines all possible Lipschitz constants without constructing a specific one. Although for the sake of simplicity only the GSA for a 1-dimensional compact interval is considered below, but it should be noted that its multivariate extension based on the use of Peano-type space-filling curves to reduce dimensionality can be found in (Strongin and Sergeyev, 2000, Chapter 8).

Algorithm 1 *The algorithm has an input $r > 1$. The function to be minimised is $Q : [a, b] \to \mathbb{R}$ with a Lipschitz constant $L > 0$.*

Initialisation: $u_0 := a$ and $u_1 := b$

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

The choice of u_{k+1} for $k \geq 1$ is determined by the following steps:

(i) *Renumber the points $u_0, \ldots, u_k$ of the previous trials by superscripts in increasing order, i.e.*

$$a = u^0 < u^1 < \ldots < u^k = b,$$

alongside the corresponding outputs $y^i := Q(u^i)$.

(ii) *Set*

$$M := \max_{1 \leq i \leq k} \left| \frac{y^i - y^{i-1}}{u^i - u^{i-1}} \right|. \quad (9)$$

(iii) *Let*

$$m := \begin{cases} 1, & M = 0; \\ rM, & M > 0, \end{cases} \quad (10)$$

where $r > 1$ is the input to the algorithm.

(iv) *For each interval $[u^{i-1}, u^i]$, $1 \leq i \leq k$, calculate the characteristic value*

$$R(i) := m(u^i - u^{i-1}) + \frac{(y^i - y^{i-1})^2}{m(u^i - u^{i-1})} - 2(y^i + y^{i-1}). \quad (11)$$

(v) *Let $R(s) := \max_{1 \leq i \leq k} R(i)$. If there are more than one solution, then the minimal integer is accepted as s.*

(vi) *Define*

$$u_{k+1} := \frac{u^s + u^{s-1}}{2} - \frac{y^s - y^{s-1}}{2m}$$

as the next trial point.

(vii) *Suppose at some iteration N the length d of the interval on which there lies a trial point $\hat{u}$ satisfying $\hat{y}_N = Q(\hat{u})$ (cf. (6)) is such that $Ld \leq \delta$, then all subsequent algorithm's outputs are set to be $\hat{u}$, i.e. $u_{k+j} := \hat{u}$ for $j = 1, 2, \ldots$.*

All of the steps except the last in Algorithm 1 are taken from (Strongin and Sergeyev, 2000, Section 3.1). Step (vii) is included additionally as a termination criterion. It is the same as that for the modified DIRECT method in Khong et al. (2013), which serves the purpose of asymptotically driving the steady-state behaviour of the plant to a neighbourhood of an extremum, without eventually leaving it; see Section 3.2.

The following result shows that given any $\mu > 0$, Algorithm 1 satisfies Assumption 7 with $\nu < \min\{(\mu - \delta)/3, 1\}$, where $\delta < \mu$ is a design parameter.

Proposition 8 *Let the point $\bar{u} \in [a, b]$ be a limit point of the sequence $\{u_k\}_{k=0}^{\infty}$ generated by Algorithm 1 for minimising a $Q : [a, b] \to \mathbb{R}$ with Lipschitz constant L whose measurements are corrupted by ν-bounded noise with $\nu < 1$. Then if at some iteration, the value m from (10) satisfies*

$$m > 2L, \quad (12)$$

$Q(\bar{u}) \leq y^* + 3\nu$, where y^* denotes the global minimum point of Q.

PROOF. Suppose (12) is met at some iteration q of Algorithm 1, then note that from (9) and (10) it will be met at any subsequent iteration $k \geq q$. Denote by $j = j(k)$ the number of the interval encompassing a global minimising point u^* (i.e. $Q(u^*) = y^*$) at the step k. Suppose to the contrapositive that

$$Q(\bar{u}) > Q(u^*) + 3\nu. \quad (13)$$

Then there exists a $p > 0$ such that for any $k \geq p$

$$u_{k+1} \notin [u^{j-1}, u^j]. \quad (14)$$

Let $t = t(k)$ denote the number of the interval $[u^{t-1}, u^t]$ containing the point $\bar{u}$ at the step k, then by (11)

$$R(t(k)) \to R,$$

where $R \leq -4Q(\bar{u}) + 4\nu + 4\nu^2$. From (13), it follows that for sufficiently large k,

$$R(t(k)) < -4Q(u^*) - 8\nu + 4\nu^2. \quad (15)$$

By the hypothesis that Q is Lipschitz with constant L and $y^i \in Q(u^i) + [-\nu, \nu]$ for $i = 0, 1, \ldots,$

$$y^{j-1} - Q(u^*) \leq L(u^* - u^{j-1}) + \nu \quad \text{and}$$
$$y^j - Q(u^*) \leq L(u^j - u^*) + \nu.$$

Summing both sides and multiplying by 2 yields

$$2(y^j + y^{j-1}) - 2L(u^j - u^{j-1}) \leq 4Q(u^*) + 4\nu.$$

By (11) and (12), it follows that

$$R(j(k)) > -4Q(u^*) - 4\nu \quad \text{for all } k > \max\{p, q\}.$$

This, together with (15) and the fact that $\nu^2 < \nu$ for positive $\nu < 1$, implies that

$$R(j(k)) > R(t(k))$$

for sufficiently large k, which in view of the update rule (v) of Algorithm 1 contradicts (14). Hence, it must be true that $Q(\bar{u}) \leq Q(u^*) + 3\nu$. □

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

Remark 9 *Condition (12) can be ensured by selecting a sufficiently large input $r > 1$ to Algorithm 1; see (10).*

4 Sampled-data extremum seeking control

The main sampled-data extremum seeking framework based on Teel and Popović (2001) is detailed in this section. A convergence proof using only the assumptions on the dynamical plant and extremum seeker stated in the previous sections is provided. The section is concluded with a procedure for realising the proposed sampled-data extremum seeking control on dynamical plants.

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

$$u(t) := u_k \quad \text{for all } t \in [kT, (k+1)T) \tag{16}$$

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

$$y_k := y(kT) \quad \text{for all } k = 1, 2, \ldots. \tag{17}$$

Figure 2 shows an extremum seeking scheme based on a sampled-data control law with period T. The following lemma on dynamical systems is needed to establish the main result of this section. The proof is based on ideas from (Nešić et al., 2013b, Prop. 1), where finite-dimensional state-space systems with asymptotically stable equilibrium points are considered. Note that infinite-dimensional systems with general attractors are accommodated here.

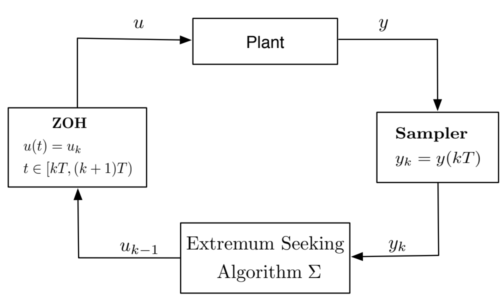

Fig. 2. Sampled-data extremum seeking control.

Lemma 10 *Given any dynamical system described in Definition 1 that satisfies Assumption 2, $\Delta > 0$, 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 (17) with y being the output of the system for the input u given by (16).

Remark 11 *The proof of Lemma 10 can be found in Khong et al. (2013). It exploits the assumption that the set of attractors is uniformly bounded with respect to all constant inputs in Ω; see (1). Alternatively, an assumption as in (Teel and Popović, 2001, Assumption 4(1)) can be made to ensure the conclusion of the lemma holds.*

The following is the main extremum seeking convergence result of this section. The feedback configuration in Figure 2 of a dynamical plant satisfying Definition 1 and Assumption 2 and an extremum seeking algorithm Σ satisfying Assumption 7, interconnected through a T-periodic sampler (17) and a synchronised zero-order hold (16), has the following convergence property:

Theorem 12 *Given any (Δ, μ) such that $\Delta, \mu > \delta$, where $\delta \geq 0$ is given in Assumption 7, there exists a sampling/waiting period $T > 0$ such that for any $\|x_0\|_{\mathcal{A}(u_0)} \leq \Delta$, $\{u_k\}_{k=0}^{\infty}$ converges to $\mathcal{C} + \mu\bar{\mathcal{B}}$, where $\mathcal{C}$ is the set of global minimisers for $Q : \Omega \subset \mathbb{R}^m \to \mathbb{R}$, the steady-state input-output map of the plant, as in Assumption 2.*

PROOF. By Assumption 7, there exists a $\nu > 0$ such that if the input to Σ, $\hat{y}$ satisfies

$$|\hat{y}_k - Q(\hat{u}_{k-1})| \leq \nu \tag{18}$$

for $k = 1, 2, \ldots$ and any $\{\hat{u}_k\}_{k=0}^{\infty} \subset \Omega$, then Σ generates an output sequence $\{u_k\}_{k=0}^{\infty}$ which converges to $\mathcal{C} + \mu\bar{\mathcal{B}}$. Furthermore, Lemma 10 ensures the existence of a sampling period $T > 0$ such that the above-mentioned sufficient condition (42) holds for any initial plant's state condition $\|x_0\|_{\mathcal{A}(u_0)} \leq \Delta$. □

Remark 13 *The above theorem holds for any dynamical plants that satisfy Definition 1 and Assumption 2. These include a broad class of general distributed-parameter nonlinear systems, such as those with periodic attractors as delineated in Section 2.2.*

4.1 Extremum seeking implementation

The following procedure summarises the implementation of the proposed extremum seeking scheme in Figure 2 to achieve a semi-global practical convergence.

Procedure 14 *Given a dynamical plant satisfying Definition 1 and Assumption 2, select an extremum seeking algorithm which has the robustness property stated in Assumption 7 with the parameter $\delta > 0$, for instance, the aforementioned Piyavskii-Shubert algorithm, DIRECT algorithm, or GSA.*

(i) *Let $u_0 \in \Omega$ be the initial trial point determined by the extremum seeking algorithm.*

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

(ii) *Select a sufficiently large region of convergence* $\Delta > \delta$ *such that* $\|x_0\|_{\mathcal{A}(u_0)} \leq \Delta$, *where* $x(0) = x_0$ *denotes the initial condition for the dynamical plant, and a sufficiently small desired accuracy of convergence to an extremum* $\mu > \delta$.
(iii) *Select a sufficiently large sampling period* $T > 0$ *corresponding to* (Δ, μ) *such that the implications of Theorem 12 hold.*
(iv) *Let* $k := 0$.
(v) *For* $t = [k, (k+1)T)$, *set the input to the plant to be* $u(t) := u_k$ *using the zero-order hold (ZOH); see (16).*
(vi) *Increment* k.
(vii) *Sample the output* y *of the plant at time* $t = kT$, *i.e.* $y_k := y(kT)$; *see (17).*
(viii) *Determine the next output of the extremum seeking controller* u_k *based on past inputs and output measurements* u_i *and* y_{i+1}, *for* $i = 0, 1, \ldots, k-1$; *see Section 3 for examples.*
(ix) *Loop from* (v). $\diamondsuit$

Remark 15 *Given the need to widely probe the input domain, practical implementations of the Piyavskii-Shubert, DIRECT and GSA extremum seekers to dynamic plants are likely to involve algorithm deployment during a calibration phase, that may be periodically repeated if the plant is known to be slowly time-varying.*

5 Asymptotically stable convergence

It is shown here that when the extremum seeking algorithms Σ satisfy a stronger convergence and robustness property than that given in Assumption 7, specifically, asymptotic stability and multi-step consistency, it is possible to establish semi-global practical asymptotic stability for the feedback system in Figure 2. This result is then related with that for the unified framework developed in the previous section. Some examples of asymptotically stable optimisation algorithms are also provided.

The convergence proof in this section is established via the use of trajectory-type properties, which is more straightforward than the Lyapunov method employed in Teel and Popović (2001); Popović (2004). The merits of doing so and the comparison between these results are provided in the succeeding section.

5.1 Convergence proof

Consider the following optimisation problem:

$$y^* := \min_{u \in \Omega} Q(u). \tag{19}$$

Assumption 16 *In reference to Figure 1, the extremum seeking controller* Σ *in Figure 2, when applied to (19), satisfies the following conditions:*

(i) Σ *is time-invariant. Denote by* $\{\hat{u}_k\}_{k=0}^{\infty} \subset \Omega$ *the output sequence* Σ *generates based on input to* Σ, $\{\hat{y}_k\}_{k=1}^{\infty}$, *where* $\hat{y}_k := Q(\hat{u}_{k-1})$. Σ *is causal in the sense that the output at any time* $N \in \mathbb{N}$, *i.e.* $\hat{u}_N$, *is determined based only on* $\hat{u}_k$ *and* $\hat{y}_{k+1}$ *for* $k = 0, 1, \ldots, N-1$, *that is the past probe values to* Q *and the corresponding measurements.*
(ii) *Denote by* $\mathcal{S}(\hat{u}_0)$ *the set of all admissible output sequences of* Σ *with respect to the initial point* $\hat{u}_0$. *There exists a class-*$\mathcal{KL}$ *function* β *such that for any initial point* $\hat{u}_0 \in \Omega$, *all outputs* $\hat{u} \in \mathcal{S}(\hat{u}_0)$ *satisfy for some* $\delta \geq 0$

$$\|\hat{u}_k(\hat{u}_0)\|_{\mathcal{C}} \leq \beta(\|\hat{u}_0\|_{\mathcal{C}}, k) + \delta \quad \forall k \geq 0. \tag{20}$$

(iii) *Let* $y_k := Q(u_{k-1}) + w_k$, *where* $w_k \in \mathbb{R}$. *Denote by* $\{u_k\}_{k=0}^{\infty}$ *the output sequence* Σ *generates based on input* $\{y_k\}_{k=1}^{\infty}$. *The pair* (u, y) *is multi-step consistent/close (Nešić et al., 1999) with* $(\hat{u}, \hat{y})$, *in the sense that for any positive* (Δ, η) *and* $N \in \mathbb{N}$, *there exists a* $\nu > 0$ *such that if* $\|u_0\|_{\mathcal{C}} \leq \Delta$ *and* $|w_k| \leq \nu$ *for* $k = 1, \ldots, N$, *then there exists a* $\hat{u} \in \mathcal{S}(u_0)$ *satisfying*

$$\|u_k - \hat{u}_k\|_2 \leq \eta \quad \text{for } k = 0, 1, \ldots, N.$$

Remark 17 *The set of outputs* $\mathcal{S}(\hat{u}_0)$ *in Assumption 16*(ii) *arises, for example, from modelling the optimisation algorithm with a difference inclusion involving a set-value 'state-update' map* F *by* $\hat{u}^+ \in F(\hat{u}, G(\hat{u}))$; *see Kellet and Teel (2005) and Section 6. In the simplest case,* $\mathcal{S}(\hat{u}_0)$ *is a singleton, i.e. there is only one possible output sequence given a fixed initial condition. For example, one modelled by a difference equation* $\hat{u}^+ = F(\hat{u}, G(\hat{u}))$.

Remark 18 *The notion of stability has not been employed in Assumption 7 of Section 3. Unlike the gradient and Newton methods to be considered later in this section, sampling algorithms like Piyavskii-Shubert, DIRECT, and GSA always initialise at a* fixed *starting point given a domain of search. As such, the usual definition of stability or attractivity (Khalil, 2002) does not apply to their outputs. Instead, what is contained in Assumption 7 is a robustness property to generate a convergent sequence in the face of perturbed measurements. In fact, it can be seen that the outputs produced by the Piyavskii-Shubert (Shubert, 1972; Nešić et al., 2013b), DIRECT (Jones et al., 1993; Khong et al., 2013), and GSA (Strongin and Sergeyev, 2000) do not satisfy (20) in general. In particular, suppose the initialisation point of these algorithms fall within* $\mathcal{C}$, *the set of global minimisers of a static objective function* Q *with only a finite number of local minima on its domain* Ω. *Observe that the updates (8) of the Piyavskii-Shubert method and GSA dictate several samples outside* $\mathcal{C} \subset \Omega$ *to be taken. The same can be said for DIRECT, due to the fact that it is set up to sample a dense subset of the whole input space*

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

Ω. The termination criterion proposed in (Khong et al., 2013) is dependent on the size of the hyper-rectangle, in which the sample that gives rise to the lowest function's output is taken, lying below a certain threshold.

The following result shows asymptotic stability of the extremum seeking scheme satisfying Assumption 16. The closed-loop system depicted in Figure 2, consisting of a dynamical plant satisfying Definition 1 and Assumption 2, T-periodic sampler (17), zero-order hold (16), and an extremum seeking algorithm satisfying Assumption 16, is asymptotically stable in the following sense:

Theorem 19 *Given any (Δ, μ) such that $\Delta, \mu > \delta$, where $\delta \geq 0$ is described in Assumption 16(ii), there exist a sampling/waiting period $T > 0$ and a $\bar{\beta} \in \mathcal{KL}$ such that for any $\|x_0\|_{\mathcal{A}(u_0)} \leq \Delta$ and $\|u_0\|_{\mathcal{C}} \leq \Delta$,*

$$\|u_k\|_{\mathcal{C}} \leq \bar{\beta}(\|u_0\|_{\mathcal{C}}, k) + \mu \qquad (21)$$

for all $k = 0, 1, \ldots$, where $\mathcal{C}$ is the set of global minimisers for $Q : \Omega \subset \mathbb{R}^m \to \mathbb{R}$, the steady-state map of the plant, as in Assumption 2.

PROOF. The proof mimics aspects of (Nešić et al., 1999, Thm. 1). Suppose β is the class-$\mathcal{KL}$ function given in Assumption 16(ii). Let $\eta > 0$ be such that

$$2\eta \leq \Delta - \delta \qquad (22)$$

and

$$\beta(2\eta + \delta, 0) \leq (\mu - \delta)/2. \qquad (23)$$

Note that η exists since $\beta(\cdot, 0) \in \mathcal{K}$ and $\beta(s, 0) \geq s - \delta$, where the last inequality holds by setting $k = 0$ in (20). By the same inequality, from (23) that $2\eta \leq (\mu - \delta)/2$, whereby

$$\eta \leq (\mu - \delta)/4. \qquad (24)$$

Now let $N \in \mathbb{N}$ be such that

$$\beta(\Delta, N) \leq \eta, \qquad (25)$$

which exists by virtue of the fact that $\beta(\Delta, \cdot)$ is decreasing to zero. With respect to the pair (Δ, η), Assumption 16(iii) guarantees the existence of a $\nu > 0$ and a $\hat{u} \in \mathcal{S}(u_0)$ such that if $\|u_0\|_{\mathcal{C}} \leq \Delta$,

$$\|u_k - \hat{u}_k\|_2 \leq \eta \quad \text{for } k = 0, 1, \ldots, N, \qquad (26)$$

whenever $|y_k - Q(u_{k-1})| \leq \nu$, $k = 1, 2, \ldots, N$. By Lemma 10, the latter can be obtained by selecting a sufficiently large waiting time or sampling period $T > 0$ for the extremum seeking scheme in Figure 2. As such, for $k = 0, 1, \ldots, N$,

$$\begin{aligned} \|u_k\|_{\mathcal{C}} &= \|\hat{u}_k + u_k - \hat{u}_k\|_{\mathcal{C}} \\ &\leq \|\hat{u}_k\|_{\mathcal{C}} + \|u_k - \hat{u}_k\|_2 \\ &\leq \beta(\|\hat{u}_0\|_{\mathcal{C}}, k) + \delta + \eta, \end{aligned} \qquad (27)$$

where (20) and (26) have been exploited. Also, since by (24) $\eta \leq (\mu - \delta)/4 < \mu - \delta$, by letting $\bar{\beta} := \beta$ it follows from (27) that (21) is satisfied for $k = 0, 1, \ldots, N$. Therefore, it remains to show that the same is true for $k > N$.

First note that by (27), (25), and (22),

$$\|u_N\|_{\mathcal{C}} \leq \beta(\Delta, k) + \delta + \eta \leq 2\eta + \delta \leq \Delta.$$

Therefore, using time-invariance, (26), (23), and (24), it follows that for all $k = N, N+1, \ldots, 2N$,

$$\begin{aligned} \|u_k\|_{\mathcal{C}} &= \|u_k(u_0)\|_{\mathcal{C}} \\ &= \|u_{k-N}(u_N)\|_{\mathcal{C}} \\ &\leq \|\hat{u}_{k-N}(u_N)\|_{\mathcal{C}} + \|\hat{u}_{k-N}(u_N) - u_{k-N}(u_N)\|_2 \\ &\leq \beta(\|u_N\|_{\mathcal{C}}, k - N) + \delta + \eta \\ &\leq \beta(2\eta + \delta, 0) + \delta + \eta \\ &\leq (\mu - \delta)/2 + \delta + (\mu - \delta)/4 \leq \mu. \end{aligned} \qquad (28)$$

Moreover, note that (28) implies

$$\|u_{2N}\|_{\mathcal{C}} \leq \beta(\Delta, N) + \delta + \eta \leq 2\eta + \delta \leq \Delta.$$

The claimed result then follows from an inductive argument. □

Remark 20 *Since Q is Lipschitz-continuous, Theorem 19 implies that there exists a $\hat{\mu} > 0$ corresponding to μ such that $Q(u_k) \to y^* + \hat{\mu}\bar{\mathcal{B}}$ as $k \to \infty$, where $\bar{\mathcal{B}}$ denotes the closed unit ball (interval) in $\mathbb{R}$. In other words, the output of the plant converges to a $\hat{\mu}$-neighbourhood of the global minimum y^*.*

5.2 Comparison with the general framework

The proof of Theorem 19 effectively demonstrates that any extremum seeking algorithm Σ satisfying Assumption 16 also satisfies Assumption 7. Note that Assumption 16 is expressed in terms of a property of an extremum seeking algorithm in the absence of measurement perturbation and a consistency condition. On the other hand, Assumption 7 directly states a robustness property of the algorithm in the face of bounded additive perturbation. The major difference between Theorem 19, which exploits a stronger Assumption 16, and Theorem 12, which is based on the more general Assumption 7, is that asymptotically stable convergence can be shown for the former, whereas only attractivity for the latter. It is known that asymptotic stability guarantees robustness to (other forms of) perturbations of

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

the closed-loop system (Khalil, 2002). This is not true in general for attractive but unstable systems.

5.3 Examples of extremum seeking algorithms

Two gradient-based examples of extremum seeking controllers are shown to satisfy Assumption 16. It is often the case that gradient-based extremum seeking algorithms can be serially decomposed into a derivative[3] estimator and a nonlinear programming method as shown in Figure 3. This paradigm is analogous to its *continuous-time* counterpart in Nešić et al. (2010, 2012), where the singular perturbation technique and time-scale separation are used to establish convergence of the extremum seeking scheme therein.

Suppose the map $Q : \Omega \subset \mathbb{R}^m \to \mathbb{R}$ is N times continuously differentiable, define

$$\mathcal{D}_Q^N(z) := \begin{bmatrix} Q(z) \\ \mathcal{D}_Q^{1,0,\ldots,0}(z) \\ \vdots \\ \mathcal{D}_Q^{N,\ldots,N}(z) \end{bmatrix},$$

where $\mathcal{D}_Q^{i_1,\ldots,i_m}(z) := \frac{\partial^{i_1+\cdots i_m} Q(z)}{\partial z_1^{i_1} \ldots \partial z_m^{i_m}}$ for some $z \in \Omega \subset \mathbb{R}^m$ and $\{z_1, \ldots, z_m\}$ denotes a set of basis vectors for $\mathbb{R}^m$.

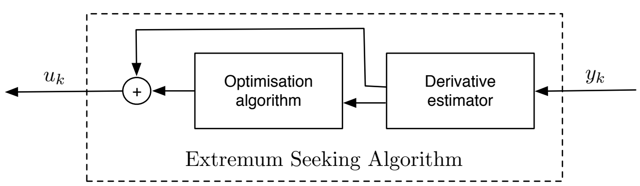

Fig. 3. A gradient-based extremum seeking controller paradigm.

Procedure 21 *Let the initial output of the extremum seeking controller be u_0. As determined by the derivative estimator, the following length-p sequence of step commands, spaced T seconds apart, can be used to probe the dynamical plant along the directions given by the basis vectors $z_1, \ldots, z_m$:*

$$(u_0 + d_1(u_0), \ldots, u_0 + d_p(u_0)), \tag{29}$$

where $d_i : \Omega \to \mathbb{R}^m$ denote the dither signals. The corresponding outputs of the plant are then sampled and collected by the derivative estimator to numerically approximate the first N-order partial derivatives of the steady-state map Q at u_0, i.e. $\mathcal{D}_Q^1(u_0), \ldots, \mathcal{D}_Q^N(u_0)$, as needed by the optimisation algorithm. This can be achieved, for example, by using the Euler methods, trapezoidal rule, or the more sophisticated Runge-Kutta methods (Press et al., 2007); see also Zong and Zhang (2009) for differential quadrature methods. Exploiting this information, the optimisation algorithm can then update its control command to u_1, and the series of steps described above repeats. In particular, $u_0 + d_p(u_0)$ and $u_1 + d_1(u_1)$ are spaced T seconds apart. ◇

[3] The first or higher derivatives of an objective function.

Two of the most well-known methods (Boyd and Vandenberghe, 2004; Polak, 1997) in operations research are (i) the *gradient descent method*:

$$u_{k+1} = u_k - \lambda_k \nabla Q(u_k),$$

where λ_i denotes the step size which can be computed by, say, the Armijo method (Polak, 1997, Alg. 1.3.3) and (ii) the *Newton's method*:

$$u_{k+1} = u_k - \nabla^2 Q(u_k)^{-1} \nabla Q(u_k),$$

where $\nabla Q(\cdot)$ and $\nabla^2 Q(\cdot)$ denote, respectively, the Jacobian and Hessian of Q. It can be readily seen that the gradient and Newton methods satisfy the time-invariance and causality Assumption 16(i). The following result can be found in Polak (1997); Boyd and Vandenberghe (2004).

Proposition 22 *Suppose $Q : \Omega \to \mathbb{R}$ is twice Lipschitz continuously differentiable and strictly convex on $\mathcal{S} \subset \Omega$, whereby there exist $m, M \in \mathbb{R}$ such that*

$$mI \leq \nabla^2 Q(u) \leq MI \quad \textit{for all } u \in \mathcal{S}.$$

Furthermore, suppose there exists a minimiser $u^* \in \mathcal{S}$ such that $\nabla Q(u^*) = 0$. Let $\{u_k\}_{k=0}^{\infty}$ be the sequence generated by the gradient or Newton method when applied to minimising Q. Then there exists a class-$\mathcal{KL}$ function β such that for any $u_0 \in \mathcal{S}$,

$$\|u_k - u^*\|_2 \leq \beta(\|u_0 - u^*\|_2, k) \quad \forall k \geq 0. \tag{30}$$

Note that the rate of convergence for the gradient descent method is linear while that for Newton is quadratic, at least within a sufficiently small neighbourhood of the minimiser.

By the converse Lyapunov theorem (Vidyasagar, 2002; Kellet, 2002), there exists a continuously differentiable positive definite Lyapunov function $V : \mathbb{Z}_+ \times \Omega \to \mathbb{R}$ such that

$$V(k+1, u_{k+1}) - V(k, u_k) \leq -\alpha(\|u_k\|_{\mathcal{C}}),$$

for some $\alpha \in \mathcal{K}$, where $\mathcal{C} := \{u^*\}$.

Suppose the use of the derivative estimates (instead of their precise values) in Figure 3 introduces a bounded

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

additive error term in the update of the gradient and Newton methods:

$$u_{k+1} = u_k - \lambda_k \nabla Q(u_k) + e_1(k, u_k) \quad \text{and}$$
$$u_{k+1} = u_k - \nabla^2 Q(u_k)^{-1} \nabla Q(u_k) + e_2(k, u_k), \tag{31}$$

where

$$\|e_1(k, u_k)\|_2 \leq l_1 + q_1 \alpha(\|u_k\|_{\mathcal{C}}) \quad \text{and}$$
$$\|e_2(k, u_k)\|_2 \leq l_2 + q_2 \alpha(\|u_k\|_{\mathcal{C}}), \tag{32}$$

for some $l_1, l_2, q_1, q_2 \geq 0$. It follows from the non-vanishing perturbation results for discrete-time systems in Cruz-Hernández et al. (1999) that the gradient/Newton-based extremum seeking controller in Figure 3 satisfies the ultimately bounded asymptotic stability Assumption 16(ii) for some $\delta > 0$, which is a function of l_1 and q_1 in the case of the gradient method and l_2 and q_2 in the case of Newton. The following is taken from (Teel and Popović, 2001, Assumption 4).

Assumption 23 *There exists an $\alpha_d \in \mathcal{K}$ and $c > 0$ such that the dither signals in (29) satisfy for each $i = 1, \ldots, p$,*

$$\|d_i(u)\|_2 \leq \alpha_d(\|u\|_{\mathcal{C}}) + c.$$

When Assumption 23 holds with $c = 0$, it follows that the step size used in estimating the derivatives converges to zero as u_k tends to the minimising set $\mathcal{C}$. This implies by the definition of differentiation that the magnitudes of the error terms e_1 and e_2 in (32) tend to zero as $k \to \infty$, i.e. $l_1 = l_2 = 0$. In other words, the perturbations are vanishing and the extremum seeking controller is asymptotically stable as in Assumption 16(ii) with $\delta = 0$ (Cruz-Hernández et al., 1999).

Consider now the case of imprecise sampling:

$$y_k := Q(u_{k-1}) + w_k, \tag{33}$$

where $|w_k| \leq \nu$ for $k = 1, 2, \ldots$ as in Assumption 16(iii). The reference Holoborodko (2008) proposes several high-frequency-noise-robust differentiators.

The following result is in order.

Proposition 24 *Suppose a gradient or Newton based extremum seeking controller depicted in Figure 3 is interconnected in feedback with a dynamical plant satisfying Assumption 2 through T-periodic sampler (17) and zero-order hold (16) as in Figure 2. Then given any positive pair (Δ, δ), there exist a $c > 0$, a sampling/waiting period $T > 0$, and a $\beta \in \mathcal{KL}$ such that if the dither signals satisfy Assumption 23 with the given c, for any $\|u_0\|_{\mathcal{C}} \leq \Delta$,*

$$\|u_k\|_{\mathcal{C}} \leq \beta(\|u_0\|_{\mathcal{C}}, k) + \delta \quad \forall k \geq 0,$$

where $\mathcal{C} := \{u^*\}$ and u^* is the global minimiser for $Q : \Omega \subset \mathbb{R}^m \to \mathbb{R}$, the steady-state input-output map of the plant which satisfies the assumptions in the statement of Proposition 22.

PROOF. Let $\Delta, \delta > 0$ be given and $l_1, q_1, l_2, q_2 > 0$ in (32) be such that the equilibrium point u^* of (31) is ultimately δ-bounded globally asymptotically stable (Cruz-Hernández et al., 1999). There exist sufficiently small $c > 0$ and $\nu > 0$ (a bound on the perturbations of sampled output measurements of the plant (33)) such that if the dithers satisfy Assumption 23 with this parameter c, the error terms arising from imprecise plant's output measurements satisfy inequalities (32) with the selected l_1, q_1, l_2, and q_2. Such a bound ν can be guaranteed by invoking Lemma 10, which involves employing a sufficiently long waiting/sampling period $T > 0$. □

The above proposition is purely one of a qualitative nature because designing the dither signals to satisfy Assumption 23 would not always be possible unless the minimising set $\mathcal{C}$ were known in advance, thereby eliminating the need of extremum seeking in the first place. In general, however, the perturbations in (32) are persistent and ultimately bounded asymptotic stability of the extremum seeking algorithms can be concluded.

Finally, it is remarked that the extremum seeking algorithm in Figure 3 based on the gradient or Newton method satisfies the multi-step consistency in Assumption 16(iii) follows from the fact that the right-hand sides of (31) are Lipschitz continuous in the error terms and Lemma 28 of the next section. Putting the results in this subsection, it can be concluded that the gradient and Newton methods satisfy Assumption 16.

6 Relation to the work by Teel and Popović

This section discusses the link between this paper and the predecessor work by Teel and Popović (Teel and Popović, 2001; Popović, 2004). The stability and consistency assumptions used are clarified and the differences and similarities/links identified.

6.1 Consistency

This subsection demonstrates that the assumptions made in Teel and Popović (2001) automatically deliver the multi-step consistency condition of Assumption 16(iii). The following difference inclusion form describing the extremum seeking/optimisation algorithm Σ is assumed in Teel and Popović (2001):

$$u^+ \in F(u, G(u)), \tag{34}$$

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

where F is an upper semi-continuous (cf. Definition 26) set-valued map (the update u^+ can be any element of the set) and G is a function that carries information regarding the estimate of the gradient of Q around u. In particular, F maps from $\mathbb{R}^m \times \mathbb{R}^p$ to subsets of $\mathbb{R}^m$,

$$G(u_k) := \begin{bmatrix} Q(u_k + d_1(u_k)) \\ \vdots \\ Q(u_k + d_p(u_k)) \end{bmatrix},$$

and $d_i : \Omega \to \mathbb{R}$, $i = 1, \ldots, p$ are dither/perturbation functions. For each $u \in \Omega$, the set $F(u, G(u))$ is nonempty and compact. See Teel (2000) for a class of Lyapunov-based nonsmooth optimisation algorithms of the form described above which employ the notion of Clarke generalised gradient. The algorithms iterate by issuing commands in the same way delineated in Procedure 21. Note that (34) is time-invariant and causal as per Assumption 16(i).

Remark 25 *Extremum seeking algorithms modelled by (34) are not a strict subset of those depicted in Figure 3. In particular, some algorithms may not admit a form explicitly separable into a derivative estimator and an optimisation subblock.*

Definition 26 *$F(u, G(u))$ is said to be an upper semi-continuous function of $u \in \Omega$ if for every $u_a \in \Omega$ and $\epsilon > 0$, there exists $\delta > 0$ such that for all $u_b \in \Omega$,*

$$\|u_a - u_b\|_2 \leq \delta \implies F(u_b, G(u_b)) \subseteq F(u_a, G(u_a)) + \epsilon\mathcal{B},$$

where $\mathcal{B}$ denotes the open unit ball in $\mathbb{R}^m$.

The following one-step consistency (i.e. closeness of solution over the next time step) assumption is taken from (Teel and Popović, 2001, Assumption 4(4)) or (Popović, 2004, Assumption 2.12(5)).

Assumption 27 *Given any $u_H \in F(u, H)$, let u_G be its closest point in the set $F(u, G(u))$, i.e.*

$$u_G := \operatorname*{arg\,min}_{u \in F(u,G(u))} \|u - u_H\|_2.$$

Then for all $\Delta > 0$, there exists $L_F > 0$ such that if $\|u\|_{\mathcal{C}} \leq \Delta$ and $\|H\|_2 \leq \Delta$, then

$$\|u_H - u_G\|_2 = \|u_H\|_{F(u,G(u))} \leq L_F \|H - G(u)\|_2.$$

The next result demonstrates that given upper semi-continuity of the set-valued function F, single-step consistency in Teel and Popović (2001) implies multi-step consistency used in this paper.

Lemma 28 *Given a difference inclusion $\hat{u}^+ \in F(\hat{u}, G(\hat{u}))$ and its perturbed form $u^+ \in F(u, H(u))$, where $\hat{u}_0 = u_0$,*

$$H(u_k) := G(u_k) + W(u_k) = G(u_k) + [w_1(u_k) \cdots w_p(u_k)]^T$$

with

$$|w_i(u)| \leq \nu \text{ for all } i = 1, \ldots, p,\ u \in \Omega, \tag{35}$$

and $F(\cdot, G(\cdot))$ is upper semi-continuous, one-step consistency in Assumption 27 implies multi-step consistency in Assumption 16(iii).

PROOF. Let $\eta > 0$ be given and $\Delta > 0$ be sufficiently large so that for all $k = 0, 1, \ldots, N$, $\|u_k\|_{\mathcal{C}} \leq \Delta$ and $\|H(u_k)\|_2 \leq \Delta$. Exploiting upper semicontinuity of F and the compactness of Ω, let $\delta > 0$ be such that for any $u_a, u_b \in \Omega$,

$$\|u_a - u_b\|_2 \leq \delta \Longrightarrow F(u_b, G(u_b)) \subseteq F(u_a, G(u_a)) + \frac{\eta}{2N}\mathcal{B}. \tag{36}$$

Now select a sufficiently small perturbation bound ν in (35) such that $L_F \|H(u_k) - G(u_k)\|_2 \leq \min\left\{\frac{\eta}{2N}, \frac{\delta}{2N}\right\}$. Consequently, by Assumption 27,

$$\|u_{k+1}\|_{F(u_k, G(u_k))} \leq \min\left\{\frac{\eta}{2N}, \frac{\delta}{2N}\right\},\ k = 0, 1, \ldots, N-1. \tag{37}$$

Let $\hat{u}_0 := u_0$, observe that (37) implies

$$\|u_1\|_{F(\hat{u}_0, G(\hat{u}_0))} \leq \min\left\{\frac{\eta}{2N}, \frac{\delta}{2N}\right\} \leq \delta \tag{38}$$

and

$$\|u_2\|_{F(u_1, G(u_1))} \leq \min\left\{\frac{\eta}{2N}, \frac{\delta}{2N}\right\}. \tag{39}$$

Let

$$\hat{u}_1 := \operatorname*{arg\,min}_{y \in F(\hat{u}_0, G(\hat{u}_0))} \|u_1 - y\|_2,$$

so that $\|u_1 - \hat{u}_1\|_2 \leq \delta$ from (38). Using this $\hat{u}_1$, application of (36) to (39) then yields

$$\|u_2\|_{F(\hat{u}_1, G(\hat{u}_1))} \leq 2\min\left\{\frac{\eta}{2N}, \frac{\delta}{2N}\right\}.$$

By applying the argument above inductively, it follows that for $k = 0, 1, \ldots, N-1$,

$$\|u_{k+1} - \hat{u}_{k+1}\|_2 \leq (k+1)\min\left\{\frac{\eta}{2N}, \frac{\delta}{2N}\right\} \leq \eta,$$

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

where

$$\hat{u}_{k+1} := \underset{y \in F(\hat{u}_k, G(\hat{u}_k))}{\arg\min} \|u_{k+1} - y\|_2,$$

as required. □

6.2 Asymptotic stability

It is of importance to note that by the Lyapunov stability result (Kellet and Teel, 2005, Thm. 2.7), the existence of a continuous Lyapunov function V for (34) presumed in (Teel and Popović, 2001, Assumption 2) or (Popović, 2004, Assumption 2.7) implies that Assumption 16(ii) is true; see also Kellet (2002). In particular, $V : \Omega \subset \mathbb{R}^m \to \mathbb{R}_{\geq 0}$ satisfies

$$\max_{w \in F(u, G(u))} V(w) - V(u) \leq -\|u\|_{\mathcal{C}} + \delta \quad \forall u \in \Omega.$$

Albeit not explicitly stated in (Teel and Popović, 2001, Assumption 2) or (Popović, 2004, Assumption 2.7), the Lyapunov function therein should actually be bounded below and above two $\mathcal{K}_\infty$ functions α_1 and α_2 as in Kellet and Teel (2005); Kellet (2002), i.e.,

$$\alpha_1(\|u\|_{\mathcal{C}}) \leq V(u) \leq \alpha_2(\|u\|_{\mathcal{C}}).$$

This is necessary for the main extremum seeking results of Teel and Popović (Teel and Popović, 2001, Rem. 5) and (Popović, 2004, Cor. 2.15) to hold, by the stability results in Kellet and Teel (2005); Kellet (2002). Indeed, (Popović, 2004, Rem. 2.9) notes that the existence of the required Lyapunov function is ascertained by a converse Lyapunov theorem in Kellet (2002).

6.3 Discussion

The work by Teel and Popović (Teel and Popović, 2001; Popović, 2004) establishes semi-global practical asymptotic stability of the sampled-data extremum scheme in Figure 2 via the Lyapunov's second method (Khalil, 2002; Kellet and Teel, 2005; Kellet, 2002) for a particular class of algorithms modelled by the state-update difference inclusion of the form (34). The main regularity condition [4] which ensures robustness to dynamical perturbation of the steady-state map is one-step consistency of the algorithm's state, as described in Assumption 27. By contrast, a trajectory-based approach is adopted in this paper to developing the same end result, but in conjunction with consistency of output of the extremum seeking algorithm over multiple time steps (cf. Assumption 16(iii)), *without* stipulating a differential inclusion model as in (34). Interestingly, Lemma 28 shows that multi-step consistency follows from one-step consistency provided the right-hand side of (34) satisfies the upper semicontinuity property given in Definition 26. In view of this result, it can be concluded that the asymptotically stable convergence of extremum seeking in Theorem 19 is developed using no more assumptions than those made in Teel and Popović (2001); Popović (2004). To be specific, Assumption 16 is a consequence of realising the extremum seeking algorithm with the differential inclusion (34) whose right-hand side is upper semicontinuous, the Lyapunov stability results in Kellet and Teel (2005); Kellet (2002), and the one-step consistency Assumption 27.

It is the authors' belief that the trajectory-based stability proofs presented in Section 5 are more direct and straightforward than the Lyapunov methods in Teel and Popović (2001); Popović (2004). Take the well-known descent methods of gradient and Newton for example, their convergence proofs are readily established via a $\mathcal{KL}$-type trajectory-based argument; see (Boyd and Vandenberghe, 2004, Chapter 9) or (Polak, 1997, Chapter 1). As a second example, recall that taking the notion of physical energy into account, there exist many systems for which Lyapunov functions whose derivative are not strictly negative can be found. The principle of Krasovskii-Lasalle (Khalil, 2002) is often exploited to conclude asymptotic stability of such systems. To apply the extremum seeking convergence result in Teel and Popović (2001) to the above cases, one would need to appeal to the converse Lyapunov theorems in Kellet (2002) to arrive at the required Lyapunov functions for the extremum seeking algorithm. Furthermore, since the main result in Teel and Popović (2001) is expressed in terms of Lyapunov functions, the user would need to apply the Lyapunov theorems in Kellet and Teel (2005); Kellet (2002) to conclude the asymptotic stability derived in Theorem 19. The results in this paper eliminate the need for such a detour, while accommodating situations where the sought Lyapunov theorems are not readily available. This may be the case when, for instance, the extremum seeking algorithms are not realisable by a difference inclusion.

In fact, Section 3 demonstrates that pursuing convergence of extremum seeking control in this direction lends insights to a developing a more fundamental framework through which to incorporate an even larger class of optimisation algorithms. For instance, those which cannot be written down as a state-update difference inclusion (34) or do not satisfy the asymptotic stability properties specified in Assumption 16.

As a final note, it is mentioned that a recurring theme of this paper is about making use of sufficiently long waiting times to decouple the plant's dynamics from the extremum seeking algorithm. The merits of using a shorter waiting time have been investigated in Popović

[4] There are other assumptions made in Teel and Popović (2001) to facilitate the Lyapunov-function-based stability proof given the difference inclusion form (34).

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

(2004) with the goal of accelerating the rate of convergence to the extremum. This effectively leads to more dynamic interaction between the plant and extremum seeker, whereby the convergence speedup is obtained at the expense of lowered accuracy.

7 Multi-unit paradigms

Prevalent in engineering is the problem of driving a finite number of almost identical systems/units to an optimal steady-state input-output behaviour, which arises, for example, in micro-array reactors and fuel cell arrays (Esmaeil-Zadeh-Azar, 2010; Woodward et al., 2009). In the presence of more than a single unit, measurement collection for the purpose of extremum seeking can be made more efficient in an appropriate fashion via parallel computation. Figure 4 shows a generalisation of the sampled-data extremum seeking control scheme of Figure 2 to multiple units. Within this context, each of these units may be driven with a different input concurrently and the corresponding outputs collected/sampled after a chosen waiting time. This results in several function evaluations of the (perturbed) steady-state behaviour within the same sampling/waiting period.

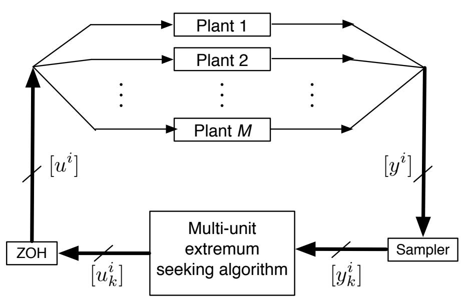

Fig. 4. Extremum seeking with multiple units

As an extension of the multi-unit framework in Khong et al. (2013), this paper considers systems or units that may exhibit different dynamics *and* steady-state input-output maps subject to a sufficiently small error bound. The performance measure of extremum seeking is dependent on an auxiliary system, which is taken to be the 'average' of all the available systems. Semi-global convergence that is practical with respect to the infinity-norm error bound on the discrepancy between units is established within the unified frameworks of Sections 4 and 5, which in the case of the latter is also asymptotically stable.

Definition 29 *Given M number of dynamical plants P_1 to P_M each satisfying Definition 2, let the average system*

$$P_a := \frac{1}{M}\sum_{i=1}^{M} P_i.$$

Denote respectively the corresponding steady-state input-output maps by $Q_1, \ldots, Q_M$ and Q_a, all mapping from $\Omega \subset \mathbb{R}^m$ into $\mathbb{R}$. Let $\mathcal{C} \subset \Omega$ be the set of global minimisers of Q_a.

Throughout, superscripts are used to label the inputs/outputs corresponding to a particular system. For instance, the input, state, and output of P_i are u^i, x^i, and y^i, respectively, for $i = 1, \ldots, M$. The T-periodic sampling and hold operations of Figure 4 are described similarly to (17) and (16) as

$$u^i(t) := u^i_k \quad \text{for all } t \in [kT, (k+1)T) \tag{40}$$

and $x^i_k := x^i(kT)$,

$$y^i_k := y^i(kT) \quad \text{for all } k = 1, 2, \ldots. \tag{41}$$

Assumption 30 *There exists a $\gamma > 0$ such that*

$$\|Q_i - Q_a\|_\infty := \sup_{u\in\Omega} |Q_i(u) - Q_a(u)| \le \gamma$$

for all $i = 1, \ldots, M$.

Remark 31 *Note that by Definition 29 the plants do not need to exhibit identical dynamics. However, it is required in Assumption 30 that their steady-state behaviours differ by no more than some bound γ in the infinity-norm sense. The average plant P_a is taken to be the nominal system. $P_1, \ldots, P_M$ can be modelled by a stochastic variable with its mean approximated by P_a, for instance.*

Based on the preceding developments, two convergence results for the sampled-data extremum seeking frameworks of Sections 4 and 5 are presented below.

7.1 The attractivity based unified framework

The following is a multi-unit expression of Assumption 7.

Assumption 32 *The extremum seeking algorithm Σ satisfies the following: There exists a δ dependent on γ in Assumption 30 such that given any $\mu > \delta$, there exists a $\nu > 0$ such that if $|\hat{y}^i_k - Q_i(\hat{u}_{k-1})| \le \nu$ for $k = 1, 2, \ldots$, $i = 1, 2, \ldots, M$, and any sequence $\{\hat{u}_k\}_{k=0}^\infty \subset \Omega$, then for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ for which $u^i_k \in \mathcal{C} + (\mu + \epsilon)\bar{\mathcal{B}}$ for all $k \ge N$. In other words, all the outputs of Σ, $\{u^i_k\}_{k=0}^\infty$ converges to a μ-neighbourhood of the set $\mathcal{C}$ of global minimisers of Q_a.*

Remark 33 *The difference between the multiple Q_i acts as additional perturbations on the steady-state input-output map Q_a. In general, the larger the γ is in Assumption 30, the larger δ in Assumption 32 will be. The latter corresponds to a lesser accuracy of convergence.*

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

The following theorem contains the main convergence result of this subsection. The feedback interconnection in Figure 4 of M dynamical plants satisfying Definition 29 and Assumption 30 and an extremum seeking algorithm Σ satisfying Assumption 32, connected through a T-periodic sampler (41) and a synchronised zero-order hold (40), has the following convergence property:

Theorem 34 *Given any (Δ, μ) such that $\Delta, \mu > \delta$, where $\delta \geq 0$ is given in Assumption 32, there exists a sampling/waiting period $T > 0$ such that for any $\|x_0^i\|_{\mathcal{A}(u_0^i)} \leq \Delta$, $\{u_k^i\}_{k=0}^{\infty}$ converges to $\mathcal{C} + \mu\bar{\mathcal{B}}$, where $\mathcal{C}$ is the set of global minimisers for $Q_a : \Omega \subset \mathbb{R}^m \to \mathbb{R}$ for $i = 1, \ldots, M$.*

PROOF. By Assumption 32, there exists a $\nu > 0$ such that if the input to Σ, $\hat{y}^i$ satisfies

$$|\hat{y}_k^i - Q_i(\hat{u}_{k-1})| \leq \nu \tag{42}$$

for $k = 1, 2, \ldots$, $i = 1, 2, \ldots, M$, and any $\{\hat{u}_k\}_{k=0}^{\infty} \subset \Omega$, then the output of Σ, i.e. $\{u_k^i\}_{k=0}^{\infty}$ converges to $\mathcal{C} + \mu\bar{\mathcal{B}}$. Furthermore, application of Lemma 10 to the i^{th} plant leads to a sampling period $T_i > 0$ such that the condition (42) holds for any initial plant's state condition $\|x_0^i\|_{\mathcal{A}(u_0^i)} \leq \Delta$. The overall sampling period for the whole feedback setup can then be taken to be $T := \max_{i=1,\ldots,M} T_i$. □

Remark 35 *An inherent issue with parallelism is that of redundancy of sample points (Strongin and Sergeyev, 2000, Chapter 5). Specifically, suppose an extremum seeking algorithm catered for parallel computations outputs $M > 1$ trial points to M dynamical systems shown in Figure 4. It is apparent that only one of these points is decided on the basis of all the previous sample values, while the rest $M - 1$ points are selected in the absence of information concerning the results of the other trials, as would otherwise be the case if a sequential or serial extremum seeking method had been employed. These points may potentially be redundant or even slow down the process of the search for an extremum. Nevertheless, this problem is in general optimisation-algorithm-specific. For example, the DIRECT method in Section 3.2 is particularly suited for parallelism since within each iteration, it normally requires more than one trial point to be collected independently of each other; see Jones et al. (1993); Khong et al. (2013). In addition, the Global Search Algorithm in Section 3.3 has also a non-redundant parallel version in (Strongin and Sergeyev, 2000, Chapter 5). Its robustness analysis is left as a future work. On the other hand, it is not clear how the Piyavskii-Shubert method in Section 3.1 can be adopted into a multi-unit framework non-redundantly.*

7.2 The asymptotic stability based unified framework

Assumption 36 *Suppose Assumption 16 is generalised to a multi-unit setting appropriately as in Assumption 32.*

Asymptotic stability of the extremum seeking scheme is shown next. The closed-loop system depicted in Figure 4, consisting of M dynamical plants satisfying Definition 29 and Assumption 30, T-periodic sampler (41), zero-order hold (40), and an extremum seeking algorithm satisfying Assumption 36, is asymptotically stable in the following sense:

Theorem 37 *Given any (Δ, μ) such that $\Delta, \mu > \delta$, where $\delta \geq 0$ is as in Assumption 36, there exist a sampling/waiting period $T > 0$ and a $\bar{\beta} \in \mathcal{KL}$ such that for any $\|x_0^i\|_{\mathcal{A}(u_0)} \leq \Delta$ and $\|u_0^i\|_{\mathcal{C}} \leq \Delta$,*

$$\|u_k^i\|_{\mathcal{C}} \leq \bar{\beta}(\|u_0^i\|_{\mathcal{C}}, k) + \mu \tag{43}$$

for all $k = 0, 1, \ldots$ and $i = 1, \ldots, M$.

PROOF. The theorem can be established using the same arguments in Theorem 19 by appropriately replacing references to the properties in Assumption 16 with those in Assumption 36, along the same lines of the proof for Theorem 34. □

Remark 38 *As demonstrated by the theorem above, another benefit of working with trajectories instead of Lyapunov functions is that generalisation of convergence results to multi-unit systems is rather straightforward. It does not involve constructing aggregate Lyapunov functions.*

For nonlinear systems whose steady-state input-output maps are multivariate, the multi-unit framework proposed above equipped with gradient-based extremum seeking controllers (cf. Section 5.3) delivers great benefits in terms of increasing the efficiency of asymptotic convergence. To be specific, in view of Procedure 21, if the number of available units M is no less than the number of dither signals p required to perform a good derivative estimate on a multidimensional space, the required measurements that

$$(u_k + d_1(u_k), \ldots, u_k + d_p(u_k)),$$

entail can be all collected within a sampling/waiting period, instead of p-multiples of it in the case of sequential extremum seeking. This is carried out by feeding $u_k^i := u_k + d_i(u_k)$ to the i^{th} plant and sampling the corresponding output y^i after T seconds, for $i = 1, \ldots, p$.

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

8 Numerical examples

Consider the following nonlinear dynamical system with periodic steady-state orbits:

$$\begin{aligned}\dot{x}_1(t) &= x_2(t);\\ \dot{x}_2(t) &= -20x_1(t) - 5x_2^3(t) + 20u(t)^2\sin(20t);\\ y_1(t) &= x_1(t),\end{aligned}$$

with initial condition $x(0) = [3, -1]^T$ and $u(t) \in [-5, 10]$ for all $t \geq 0$. Denote by Q_p the steady-state input-output map of the plant. See Section 2.2 for a general description of models of this type. The following cost function is introduced as in (4):

$$y(t) = L_\infty(y_1(t)) := \max_{\tau \in [t, t-\frac{\pi}{10})} |y_1(\tau)| \geq 0.$$

The global minimum of the steady-state map with respect to the cost function $Q := L_\infty \circ Q_p$ is located at $u = 0$, which leads to $y = 0$. To see this, note that when $u = 0$, the steady state of the system is given by $x_1 = x_2 = 0$, whereby $y_1 = 0$. Thus, $y = L_\infty(y_1) = 0$.

8.1 Single unit

The extremum seeking control scheme in Figure 2 is applied with the controller being the DIRECT optimisation method described in Section 3.2. Figure 5 shows the output of the system over time. Using a waiting time or sampling period of $T = 1$s, $L = 2$ and $\delta = 0.02$, it takes 33s to locate $u = 0.0031$ as an estimate of the global minimum.

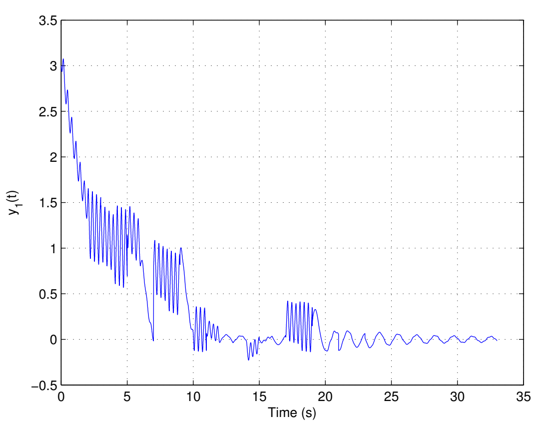

Fig. 5. Extremum seeking based on modified DIRECT

Now the extremum seeking controller based on the gradient descent method is employed with a fixed step size of 1, as described in Section 5.3. The smooth low noise differentiator with a filter length 5 from Holoborodko (2008) is used as the derivative estimator:

$$Q'(u) \approx \frac{2(Q_1 - Q_{-1}) + Q_2 - Q_{-2}}{8h}, \qquad (44)$$

where $Q_i := Q(u + ih)$ for $i = -2, -1, 1, 2$ and $h := 0.5$ denotes the estimation step size. The algorithm is terminated when an input $|u| \leq 0.05$ is found. With a sampling period $T := 1$s and an initial guess of -5, it takes 412s to locate an input $u = -0.049$; see Figure 6 for an output response of the dynamical system. The time is mainly spent on taking output measurements of the system to make good derivative estimates; (44) requires 4 points to be collected in order to produce one estimate.

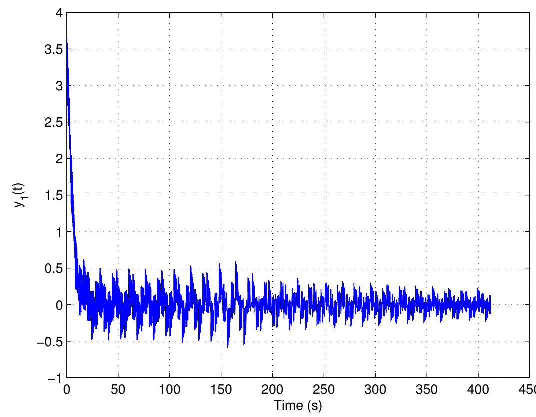

Fig. 6. Extremum seeking based on the gradient method

8.2 Double units

Suppose now there exists a second similar dynamical system at disposal:

$$\begin{aligned}\dot{x}_1(t) &= x_2(t);\\ \dot{x}_2(t) &= -20x_1(t) - 5x_2^3(t) + 20(1.01u(t))^2\sin(20t);\\ y_2(t) &= x_1(t),\end{aligned}$$

with initial condition $x(0) = [3, -1]^T$ and $u(t) \in [-5, 10]$ for all $t \geq 0$. By implementing the multi-unit extremum seeking scheme in Figure 4 with the modified DIRECT method as delineated in Remark 35, one yields the system output responses in Figure 7. In particular, it takes 13s to locate $u = 0.083$.

On the other hand, with a multi-unit extremum seeking scheme based on the gradient method with a fixed step size of 0.5, it takes 186s to find an estimate $u = -0.029$. The output responses of the systems are shown in Figure 8.

Table 1 summarises the simulation results. Note that the DIRECT method based extremum seeking controller significantly outperforms the gradient method one in

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

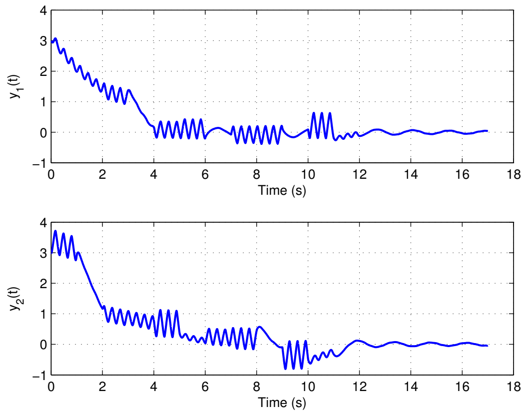

Fig. 7. Multi-unit extremum seeking based on the DIRECT method

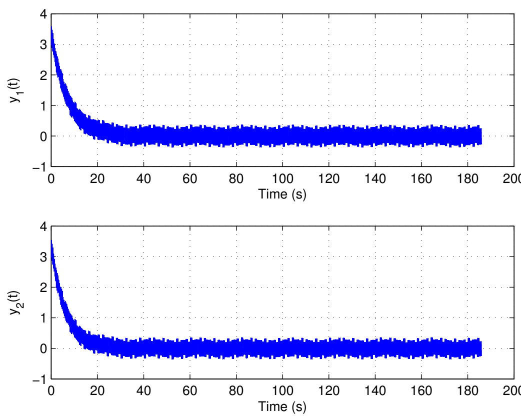

Fig. 8. Multi-unit extremum seeking based on the gradient method

terms of speed of convergence for this particular example and the parameters and derivative estimation scheme chosen. Convergence can be accelerated at the expense of accuracy by decreasing the sampling period.

Alg.	Unit	Duration	Estimate
DIRECT	Single	23s	0.031
	Multiple	13s	0.083
Gradient	Single	412s	−0.049
	Multiple	186s	−0.029

Table 1
Extremum seeking with the DIRECT and gradient methods.

9 Conclusions

Two unified periodic sampled-data frameworks for extremum seeking have been proposed for nonlinear systems with possibly infinite dimension and compact attractors. The first is based on the notion of attractivity towards an extremum of the steady-state input-output mapping of a system. This accommodates a whole range of sampling type optimisation algorithms, which are typically efficient for global, nonconvex, and nonsmooth optimisation. The second is based on asymptotic stability, for which the proof provided is simpler and uses easier-to-verify trajectory-based assumptions than those in the literature (Teel and Popović, 2001; Popović, 2004). Several gradient based methods are known to belong to the class of extremum seeking algorithms considered within this framework. Multi-unit extremum seeking is also investigated as a means to expedite the speed of convergence.

References

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

Ariyur, K. B. and Krstić, M. (2004). Slope seeking: a generalization of extremum seeking. *International Journal of Adaptive Control and Signal Processing*, 18(1):1–22.

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

Cruz-Hernández, C., Alvarez-Gallegos, J., and Castro-Linares, R. (1999). Stability of discrete nonlinear systems under nonvanishing perturbations: application to a nonlinear model-matching problem. *IMA Journal of Mathematical Control and Information*, 16:23–41.

Dower, P. M., Farrell, P., and Nešić, D. (2007). Extremum seeking control of cascaded raman optical amplifiers. *IEEE Trans. Control Syst. Technol.*, 16:296–407.

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

Ghaffari, A., Krstić, M., and Nešić, D. (2012). Multivariable Newton-based extremum seeking. *Automatica*, 48:1759–1767.

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

Haring, M., van de Wouw, N., and Nešić, D. (2012). Extremum-seeking control for nonlinear systems with periodic steady-state outputs. *Automatica*. To appear.

Holoborodko, P. (2008). Smooth noise robust differentiators. http://www.holoborodko.com/pavel/numerical-methods/numerical-derivative/smooth-low-noise-differentiators/.

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.

Kellet, C. M. (2002). *Advances in Converse and Control*

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST

Lyapunov Functions. PhD thesis, University of California, Santa Barbara.

Kellet, C. M. and Teel, A. R. (2005). On the robustness of $\mathcal{KL}$-stability for diffrence inclusions: smooth discrete-time Lyapunov functions. *SIAM J. Control Optim.*, 44(3):777–800.

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

Khong, S. Z., Nešić, D., Manzie, C., and Tan, Y. (2013). Multidimensional global extremum seeking via the DIRECT method. *Automatica*. in press.

Krstić, M. and Wang, H.-H. (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, 36(4):595–601.

Kvaternik, K. and Pavel, L. (2011). Interconnection conditions for the stability of nonlinear sampled-data extremum seeking schemes. In *Proc. 50th IEEE Conf. Decision Control*.

Luenberger, D. G. (1969). *Optimization by Vector Space Methods*. Series in Decision and Control. Wiley-Interscience.

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

Nešić, D. (2009). Extremum seeking control: convergence analysis. *Europ J. Contr.*, 3(4):331–374.

Nešić, D., Mohammadi, A., and Manzie, C. (2013a). A framework for extremum seeking control of systems with parameter uncertainties. *IEEE Trans. Autom. Contr.* In press.

Nešić, D., Nguyen, T., Tan, Y., and Manzie, C. (2013b). A non-gradient approach to global extremum seeking: an adaptation of the Shubert algorithm. *Automatica*. In press.

Nešić, D., Tan, Y., Manzie, C., Mohammadi, A., and Moase, W. H. (2012). A unifying framework for anaysis and design of extremum seeking controllers. In *Proceedings of Chinese Control and Decision Conference*.

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

Nešić, D., Teel, A. R., and Kokotović, P. V. (1999). Sufficient conditions for stabilization of sampled-data nonlinear systems via discrete-time approximations. *Systems & Control Letters*, 38:259–270.

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

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

Polak, E. (1997). *Optimization — Algorithms and Consistent Approximations*. Applied Mathematical Sciences 124. Spinger.

Popović, D. (2004). *Topics in Extremum Seeking*. PhD thesis, University of California, Santa Barbara.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, 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.

Tan, Y., Nešić, D., and Mareels, I. M. Y. (2008). On the dither choice in extremum seeking control. *Automatica*, 44(5):1446–1450.

Teel, A. R. (2000). Lyapunov methods in nonsmooth optimization, part I: Quasi-Newton algorithms for Lipschitz, regular functions and part II: persistently exciting finite differences. In *Proceedings of IEEE Conference on Decision and Control*.

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.

Vidyasagar, M. (2002). *Nonlinear Systems Analysis*. SIAM, 2nd edition.

Wang, H.-H., Krstić, M., and Bastin, G. (1999a). Optimizing bioreactors by extremum seeking. *International Journal of Adaptive Control and Signal Processing*, 13:651–699.

Wang, H.-H., Yeung, S., and Krstić, M. (1999b). Experimental application of extremum seeking on an axial-flow compressor. *IEEE Trans. Control Syst. Technol.*, 8:300–309.

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. Spinger.

Zong, Z. and Zhang, Y. (2009). *Advanced Differential Quadrature Methods*. CRC Press.

Preprint submitted to Automatica
Received June 5, 2013 00:26:27 PST