Automatica 56 (2015) 44–52

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Extremum seeking of dynamical systems via gradient descent and stochastic approximation methods☆

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

[a] Department of Automatic Control, Lund University, SE-221 00 Lund, Sweden

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

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

ARTICLE INFO

Article history:
Received 15 August 2013
Received in revised form
13 October 2014
Accepted 8 March 2015
Available online 11 April 2015

Keywords:
Extremum seeking
Infinite-dimensional nonlinear systems
Sampled-data control
Gradient descent method
Stochastic approximation methods

ABSTRACT

This paper examines the use of gradient based methods for extremum seeking control of possibly infinite-dimensional dynamic nonlinear systems with general attractors within a periodic sampled-data framework. First, discrete-time gradient descent method is considered and semi-global practical asymptotic stability with respect to an ultimate bound is shown. Next, under the more complicated setting where the sampled measurements of the plant's output are corrupted by an additive noise, three basic stochastic approximation methods are analysed; namely finite-difference, random directions, and simultaneous perturbation. Semi-global convergence to an optimum with probability one is established. A tuning parameter within the sampled-data framework is the period of the synchronised sampler and hold device, which is also the waiting time during which the system dynamics settle to within a controllable neighbourhood of the steady-state input–output behaviour.

1. Introduction

Extremum seeking locates via online computations an optimal operating regime of the steady-state input–output map of a dynamical system without explicit knowledge of a model (Ariyur & Krstić, 2003; Zhang & Ordóñez, 2011). Two categories of extremum seeking controllers can be found in the literature. The first of which is continuous-time controllers which exploit dither/excitation signals to probe the local behaviour of the system to be optimised and continuously transition the system input to one that results in an optimum. See Ariyur and Krstić (2003), Krstić and Wang (2000) and Tan, Nešić, and Mareels (2006) for such methods that utilise periodic dithers and Liu and Krstić (2012), Manzie and Krstić (2009) for stochastic dithers. The convergence proofs of the former rely on averaging and singular perturbation techniques (Khalil, 2002; Teel, Moreau, & Nešić, 2003), while the latter on stochastic averaging (Liu & Krstić, 2012). On the contrary, discrete-time extremum seeking controllers based on nonlinear programming methods are examined in Teel and Popović (2001) within a sampled-data framework. The convergence proof therein is established using Lyapunov arguments.

An alternative and more direct proof for convergence to an extremum in a sampled-data framework is given in Khong, Nešić, Tan, and Manzie (2013) using trajectory-based techniques. In the same paper, the sampled-data framework of extremum seeking is further examined to accommodate global nonconvex optimisation methods, such as those described in Strongin and Sergeyev (2000). These results demonstrate that a wide range of optimisation algorithms in the literature can be applied to extremum seeking of dynamic plants. Making use of the results in Khong, Nešić, Tan et al. (2013), deterministic gradient descent based extremum seeking control is reviewed in this paper. Furthermore, stochastic gradient descent (a.k.a. stochastic approximation) methods are accommodated for extremum seeking in a way that is robust against measurement errors.

Stochastic approximation methods (Kushner & Clark, 1978; Kushner & Yin, 2003; Spall, 2003) are a family of well-studied iterative gradient-based optimisation algorithms that find applications in a broad range of areas, such as adaptive control and neural networks (Bertsekas & Tsitsiklis, 1996). In contrast to the standard

☆ This work was supported by the Swedish Research Council through the LCCC Linnaeus centre and the Australian Research Council (DP120101144). The material in this paper was presented at the 9th Asian Control Conference, June 23–26, 2013, Istanbul, Turkey. This paper was recommended for publication in revised form by Associate Editor Raul Ordóñez under the direction of Editor Miroslav Krstic.

E-mail addresses: seizhen@control.lth.se (S.Z. Khong), yingt@unimelb.edu.au (Y. Tan), manziec@unimelb.edu.au (C. Manzie), dnesic@unimelb.edu.au (D. Nešić).

[1] Tel.: +46 462220362; fax: +46 46138118.

http://dx.doi.org/10.1016/j.automatica.2015.03.018
0005-1098/

optimisation algorithms such as the steepest descent or Newton methods (Boyd & Vandenberghe, 2004) which exploit direct gradient information, stochastic approximation methods operate based on *approximation* to the gradient constructed from noisy measurements of the objective/cost function. For the former, knowledge of the underlying system input–output relationships are often needed to calculate the gradient using for example, the chain rule. This is not necessary for stochastic approximation, making it well-suited for non-model based extremum seeking control.

This paper adapts within a periodic sampled-data framework three discrete-time multivariate stochastic approximation algorithms for extremum seeking control of dynamical systems which can be of infinite dimension and contain general attractors. Namely, Kiefer–Wolfowitz–Blum's Finite Difference Stochastic Approximation (FDSA) (Blum, 1954; Kiefer & Wolfowitz, 1952), Random Directions Stochastic Approximation (RDSA) (Kushner & Clark, 1978), and Simultaneous Perturbation Stochastic Approximation (SPSA) (Spall, 1992, 2003). It is shown that there exists a sufficiently long sampling period under which semi-global convergence with probability one to an extremum of the steady-state input–output relation can be achieved. This stands in comparison with the gradient descent method based extremum seeking control under ideal noise-free sample measurements, for which semi-global practical ultimately bounded *asymptotic* stability can be established. Note that the existence of Lyapunov functions satisfying the conditions in Teel and Popović (2001) is not known for the stochastic approximation methods, and hence the convergence results therein do not directly generalise to these methods.

A related work (Nusawardhana & Żak, 2004) considers an extremum seeking method based on the SPSA within a different setup (i.e. not sampled-data and has continuous plant output measurements). There, the steady-state input–output objective function is assumed to evaluate to a constant after some waiting time with respect to a constant input, and the output measurements are corrupted by noise. By contrast, this paper exploits the fact that the state trajectory of an asymptotically stable dynamical system converges to a neighbourhood of its steady-state value after the system's input is held constant for a pre-selected waiting time. Furthermore, the sampled output value is assumed to be corrupted by measurement noise. The SPSA method has also been applied to optimisation of variable cam timing engine operation in Popović, Janković, Magner, and Teel (2006), alongside several other optimisation algorithms. Azuma, Sakar, and Pappas (2012) adapts the SPSA method for *extreme source seeking* of randomly switching *static* distribution fields using a nonholonomic mobile robot. On a different note, Stanković and Stipanović (2010) considers a related problem of extremum seeking of *static* functions under noisy measurements using a discrete-time controller with sinusoidal dither signals. These works differ from the setting of the paper, where stochastic approximation methods based extremum seeking of the steady-state input–output maps of *dynamical* systems is analysed within a sampled-data framework.

The paper has the following structure. First, the next section states the properties of the nonlinear dynamical systems to which gradient descent and stochastic optimisation methods are applied. Section 3 depicts the sampled-data framework in which extremum seeking control is analysed. Subsequently, Section 4 examines the gradient descent method for extremum seeking. Stochastic optimisation methods are considered in Section 5. Illustrative simulation examples are provided in Section 6, followed by some concluding remarks in Section 7.

2. Dynamical systems

The class of nonlinear, possibly infinite-dimensional, systems with general attractors considered in this paper is introduced in this section. 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 $u : \mathbb{R}_{\geq 0} \to \mathbb{R}^m$ and $y : \mathbb{R}_{\geq 0} \to \mathbb{R}$. Given any $u \in \Omega \subset \mathbb{R}^m$ 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.

Parts of the following assumption are based on Teel and Popović (2001, Assumption 1). Remarks on each of the assumptions follow.

Assumption 2. Given a system described in Definition 1 and an open bounded set $\Omega \subset \mathbb{R}^m$, the following hold:

(i) There exists a function $\mathcal{A}$ mapping from $\mathbb{R}^m$ 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) which satisfies:
  (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,
  $$\sup_{u\in\Omega}\sup_{x\in\mathcal{A}(u)} \|x\| < \infty.$$
(ii) 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 $\|x_0\|_{\mathcal{A}(u)} \leq \Delta$.
(iii) 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 $u \in \mathbb{R}^m$ 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.
(iv) Q is thrice continuously differentiable and has bounded derivatives on Ω.
(v) The Jacobian $\nabla Q = 0$ in a nonempty, compact set $\mathcal{C} \subset \mathbb{R}^m$, i.e. Q achieves its minimum on $\mathcal{C}$.

Remark 3. Property (i) of Assumption 2 states that for each constant input to the system, there exists a corresponding set to which the state of the system converges. Property (ii) stipulates that the state converges asymptotically stably. Property (iii) guarantees the existence of a corresponding output and hence an input–output map Q in steady state. The last two conditions are properties of Q which are assumed for convergence of the approximate gradient optimisation methods used in this paper, and are consistent with corresponding assumptions in e.g. Spall (2003).

An example of systems satisfying the above assumption follows. Consider the following dynamical system (Ariyur & Krstić, 2003; Tan et al., 2006):

$$\begin{aligned} &\dot{x} = f(x, u) \quad x(0) = x_0; \\ &y = h(x), \end{aligned} \tag{1}$$

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 $\ell : \mathbb{R}^m \to \mathbb{R}^n$ such that

$$f(\ell(u), u) = 0 \quad \forall u \in \mathbb{R}^m.$$

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

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

where $x(\cdot, x_0, u)$ denotes the solution to (1) with respect to the initial condition x_0 and input u.

Definition 5. Let

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

be the steady-state input–output map of system (1).

Suppose Q achieves its minimum value on a nonempty compact set $\mathcal{C} \subset \Omega$ and is continuously differentiable with bounded derivatives. It can then be verified easily that if the system (1) satisfies Assumption 4, it also satisfies Assumption 2. In particular, Assumption 2(i) and (ii) are implied by Assumption 4 with $\mathcal{A}(u) = \ell(u)$, i.e. a singleton, for every $u \in \mathbb{R}^m$. Assumption 2(iii) follows immediately from Definition 5.

In essence, a large class of nonlinear systems satisfies Assumption 2, including as a special case finite-dimensional state-space systems with equilibria of the form (1) satisfying Assumption 4 and the conditions in Assumption 2(iv)–(v) on Q. It opens up a wider class of systems to be addressed, such as those that are possibly of infinite dimension (with states living in an abstract Banach space $\mathcal{X}$) and may admit more general attractors than equilibria; see Khong, Nešić, Tan et al. (2013).

3. Sampled-data extremum seeking framework

The sampled-data extremum seeking framework of Khong, Nešić, Tan et al. (2013), Teel and Popović (2001) is detailed in this section. The gradient based methods in the succeeding sections can be applied to extremum seeking of dynamical systems defined in Section 2 within this framework.

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

$$u(t) := u_k \quad \text{for all } t \in [kT, (k+1)T) \tag{2}$$

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 $x_k := x(kT)$;

$$y_k := y(kT) \quad \text{for all } k = 1, 2, \ldots. \tag{3}$$

Fig. 1 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 results of the next sections. The proof is based on ideas from Nešić, Nguyen, Tan, and Manzie (2013, 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.

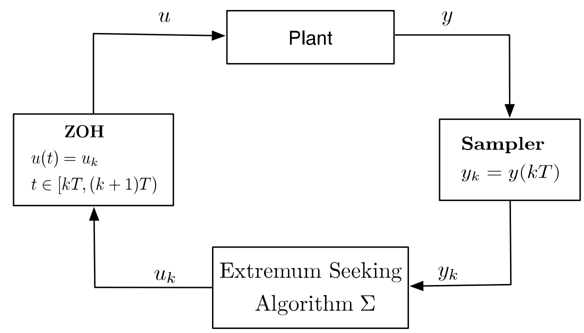


Fig. 1. Sampled-data extremum seeking control.

Lemma 6 (*Khong, Nešić, Manzie, & Tan, 2013*, *Lem. 13*). *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 $\{u_k\}_{k=0}^{\infty} \subset \mathbb{R}^m$ and $\|x_0\|_{\mathcal{A}(u_0)} \le \Delta$,*

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

where x_0 denotes the initial condition of the system and y_k is as in (3) with y being the output of the system for the input u given by (2).

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

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

where $\mathcal{A}_{\max} := \sup_{u \in \Omega} \sup_{x \in \mathcal{A}(u)} \|x\|$, which is finite by Assumption 2(i). Choose $\epsilon_1 > 0$ and $\epsilon_2 > 0$ so that

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

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

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

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

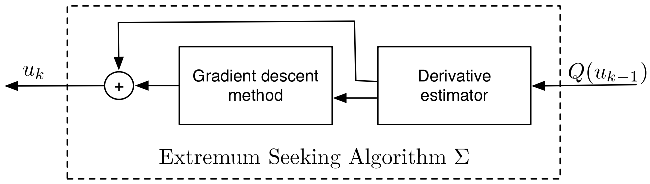


Fig. 2. A gradient-based extremum seeking controller paradigm.

4. Gradient descent method

This section adapts the gradient descent method for extremum seeking control within the sampled-data setting introduced in the previous section. The method may be considered as a special case under the unified framework proposed in the paper (Khong, Nešić, Tan et al., 2013), of which several results are utilised here.

Consider the following optimisation problem:

$$y^* := \min_{u\in\Omega} Q(u), \tag{4}$$

where $Q : \mathbb{R}^m \to \mathbb{R}$ and $\Omega \subset \mathbb{R}^m$. Assuming that Q is differentiable, one of the most used methods in operations research is the *gradient descent method*:

$$\theta_{j+1} = \theta_j - \lambda_j \nabla Q(\theta_j), \tag{5}$$

where λ_j denotes the step size at time j and $\nabla Q(\cdot)$ the Jacobian of Q. The following result is standard in convex optimisation (Boyd & Vandenberghe, 2004; Polak, 1997).

Proposition 7. *Suppose $Q : \mathbb{R}^m \to \mathbb{R}$ is twice continuously differentiable with bounded derivatives and strictly convex on $\Omega \subset \mathbb{R}^m$, whereby there exist $\underline{M}, M > 0$ such that*

$$\underline{M}I \le \nabla^2 Q(u) \le MI \quad \text{for all } u \in \Omega,$$

where $\nabla^2 Q(\cdot)$ denotes the Hessian of Q. Furthermore, suppose there exists a minimiser $\theta^* \in \Omega$ such that $\nabla Q(\theta^*) = 0$. Let $\mathcal{C} := \{\theta^*\}$ and $\{\theta_j\}_{j=0}^{\infty} \subset \Omega$ be the sequence generated by the gradient method with fixed step size

$$\lambda \le \frac{2}{\underline{M} + M}$$

applied to minimising Q. Then there exists a class-$\mathcal{KL}$ function β such that for any $\theta_0 \in \Omega$,

$$\begin{aligned}\|\theta_j\|_{\mathcal{C}} := \|\theta_j - \theta^*\|_2 &\le \beta(\|\theta_0\|_{\mathcal{C}}, j)\\ &= \beta(\|\theta_0 - \theta^*\|_2, j) \quad \forall j = 0, 1, \ldots.\end{aligned} \tag{6}$$

In particular, $\beta(\|\theta_0 - \theta^*\|_2, j)$ can be taken to be $c^{\frac{j}{2}}\|\theta_0 - \theta^*\|_2$ with $c := 1 - \lambda\frac{2\underline{M}M}{\underline{M}+M}$. This implies linear rate of convergence $O(1/j)$.

Proposition 7 states the convergence conditions for the gradient descent method as an extremum seeking algorithm when exact values of $\nabla Q(\theta_j)$ are known for implementing (8). In practice, the Jacobian $\nabla Q(\theta_j)$ needs to be estimated from several past measurements. This can be achieved by using the Euler methods, trapezoidal method, or the more sophisticated Runge–Kutta methods (Press, Teukolsky, Vetterling, & Flannery, 2007); see Fig. 2. Henceforth, the gradient descent based extremum seeking algorithm is called Σ.

To be more specific, let the initial output of the extremum seeking controller be $u_0 := \theta_0$. As determined by the derivative estimator, the following length-p sequence of step commands $\{u_k\}_{k=0}^{p-1}$ can be used to probe Q along the desired directions:

$$(\theta_0 + d_1(\theta_0), \ldots, \theta_0 + d_p(\theta_0)), \tag{7}$$

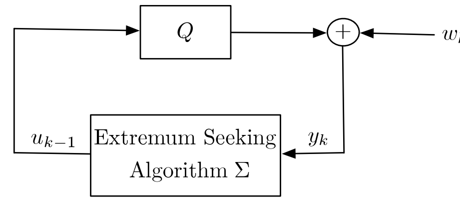


Fig. 3. Extremum seeking algorithm with noisy output measurement.

where $d_i : \Omega \to \mathbb{R}^m$, $i = 1, \ldots, p$ denote the dither signals. The corresponding outputs of Q are then collected by the derivative estimator to numerically approximate the Jacobian $\nabla Q(\theta_0)$. Exploiting this information, the optimisation algorithm can then update its control command to θ_1, and the series of steps described above repeats to generate $\{u_k\}_{k=p}^{2p-1}$.

Suppose the use of the derivative estimates introduces a bounded additive error term in the update of the gradient descent method:

$$\theta_{j+1} = \theta_j - \lambda_k \nabla Q(\theta_j) + e(\theta_j) \tag{8}$$

where $e(\cdot)$ is continuous and satisfies

$$\|e(\theta_j)\|_2 \le l + q\alpha(\|\theta_j\|_{\mathcal{C}}) \tag{9}$$

for some $\alpha \in \mathcal{K}$ and $l, q \ge 0$ as determined by the estimation method and step size used. It follows from the non-vanishing perturbation results for discrete-time systems in Cruz-Hernández, Alvarez-Gallegos, and Castro-Linares (1999) that for sufficiently small l and q, the gradient-based extremum seeking controller in Fig. 2 satisfies the ultimately bounded asymptotic stability for some ultimate bound which is a class-$\mathcal{K}$ function of l. That is, there exist a class-$\mathcal{K}$ function α and a class-$\mathcal{KL}$ function β such that

$$\|\theta_j\|_{\mathcal{C}} \le \beta(\|\theta_0\|_{\mathcal{C}}, j) + \alpha(l) \quad \forall j = 0, 1, \ldots. \tag{10}$$

Consider now the case where output measurements of Q are corrupted by bounded perturbations as in Fig. 3. 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 $\{y_k\}_{k=1}^{\infty}$. Given θ_j, it can be seen that there exist $\alpha_{ih} \in \mathcal{K}$ for $i = 1, \ldots, m$, $h = 1, \ldots, p$ such that the update of the gradient descent method (8) becomes

$$\begin{aligned}\theta_{j+1} &= \theta_j - \lambda_j \nabla Q(\theta_j) + e(\theta_j)\\ &+ \begin{bmatrix} \alpha_{11}(|w_{jp+1}|) + \cdots + \alpha_{1p}(|w_{(j+1)p}|) \\ \vdots \\ \alpha_{m1}(|w_{jp+1}|) + \cdots + \alpha_{mp}(|w_{(j+1)p}|) \end{bmatrix}.\end{aligned} \tag{11}$$

The $\mathcal{K}$ functions model the fact that the error on the update increases with the magnitude of w_k.

Let $\{\hat{u}_k\}_{k=0}^{\infty} \subset \Omega$ denote the nominal output sequence Σ generates based on the uncorrupted input to Σ, $\{\hat{y}_k\}_{k=1}^{\infty}$ with $\hat{y}_k := Q(\hat{u}_{k-1})$. Note that the pair (u, y) is multi-step consistent/close (Nešić, Teel, & Kokotović, 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 $\|u_0\|_{\mathcal{C}} \le \Delta$ and $|w_k| \le \nu$ for $k = 1, \ldots, N$, then

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

This can be established from the fact that (u, y) is clearly one-step consistent with $(\hat{u}, \hat{y})$ by (11), the right-hand side of (8) is continuous in θ_j, and Khong, Nešić, Tan et al. (2013, Lem. 28), which demonstrates in this case that one-step consistency implies multi-step consistency.

The following result shows asymptotic stability of the gradient descent based extremum seeking scheme Σ. The closed-loop system depicted in Fig. 1, consisting of a dynamical plant satisfying Definition 1 and Assumption 2 with a steady-state input–output map Q that satisfies the conditions in Proposition 7, T-periodic sampler (3), zero-order hold (2), and a gradient descent based extremum seeking algorithm Σ with some chosen derivative estimator shown in Fig. 2, is semi-globally practically asymptotically stable with respect to an ultimate bound in the following sense:

Theorem 8. *Given any (Δ, μ) such that $\Delta, \mu > \alpha(l)$ in (10), there exist a sampling/waiting period $T > 0$ and a $\beta \in \mathcal{KL}$ such that for any initial state $\|x_0\|_{\mathcal{A}(\theta_0)} \le \Delta$ and $\|\theta_0 - \theta^*\|_2 \le \Delta$,*

$$\|\theta_j - \theta^*\|_2 \le \beta(\|\theta_0 - \theta^*\|_2, j) + \mu \tag{12}$$

for all $j = 0, 1, \ldots$, where θ^* satisfies $\nabla Q(\theta^*) = 0$.

Proof. This follows from Khong, Nešić, Tan et al. (2013, Thm. 19). In particular, the multi-step consistency and time-invariance of Σ are exploited to show asymptotically stable convergence to a μ-neighbourhood of θ^* for a sufficiently small perturbations magnitude ν, which can be guaranteed by application of Lemma 6 via the use of a sufficiently long sampling period T in the sampled-data extremum seeking control framework. □

While flexibility is present in selecting a derivative estimator when implementing the gradient descent extremum seeking controller in this section, the stochastic approximation methods to be discussed in the next section all employ variants of the one-step Euler method to generate derivative estimates. It is assumed there that the estimates are generated based on *noisy measurements*. A major difference in the conclusions is that asymptotically stable practical convergence to a neighbourhood of the minimiser is shown in this section while only attractivity towards the minimiser with probability one (w.p.1) can be established in the next. It is known that asymptotic stability guarantees robustness to different forms of perturbations on closed-loop systems (Khalil, 2002). This is not true in general for attractive but Lyapunov unstable systems.

5. Stochastic approximation

This section adapts three stochastic approximation algorithms for extremum seeking control and establishes semi-global convergence with probability one. It is divided into three subsections, respectively dedicated to the Finite Difference Stochastic Approximation (FDSA), Random Directions Stochastic Approximation (RDSA), and Simultaneous Perturbation Stochastic Approximation (SPSA). Every subsection begins by reviewing the respective stochastic approximation (SA) algorithm. The review material is largely based on Chin (1997), Kushner and Clark (1978) and Spall (2003).

All three algorithms are based around the following basic structure. Let $Q : \Omega \subset \mathbb{R}^m \to \mathbb{R}$ be a thrice continuously differentiable objective/cost function with bounded derivatives and $g(\theta) := \nabla Q(\theta)$. Stochastic approximation algorithms are iterative procedures which find a minimising θ^* of Q such that $g(\theta^*) = 0$. They take the following standard form

$$\theta_{j+1} = \theta_j - a_j g_j(\theta_j), \quad j = 0, 1, \ldots, \tag{13}$$

where $\{a_j\}_{j=0}^{\infty}$ is a positive gain sequence and $g_j(\theta_j)$ is an approximation to the gradient g of Q at θ_j. The approximation is constructed from noisy measurements of Q at appropriate points in its domain Ω.

The ensuing subsections detail three stochastic approximation methods and their convergence conditions for extremum seeking control.

5.1. Finite difference (FDSA)

Let $\{e_i\}_{i=1}^m$ be the canonical basis for $\mathbb{R}^m$, i.e. e_i is a unit vector in the direction of the ith coordinate of $\mathbb{R}^m$. The FDSA (Blum, 1954; Kiefer & Wolfowitz, 1952) utilises a finite-difference equation to approximate each entry of the gradient g. In particular, let $j \in \{0, 1, \ldots\}$ be the current iteration number and a positive scalar c_j be given, the (two-sided) FDSA takes measurements of Q at design levels $\theta_j \pm c_j e_j$:

$$\begin{aligned} y_j^{i+} &= Q(\theta_j + c_j e_i) + \epsilon_j^{i+} \\ y_j^{i-} &= Q(\theta_j - c_j e_i) + \epsilon_j^{i-}, \end{aligned} \tag{14}$$

where ϵ_j^{i+} and ϵ_j^{i-} represent measurement noise terms that satisfy

$$E\left\{\epsilon_j^{i+} - \epsilon_j^{i-} \,\middle|\, \theta_i;\ i \le j\right\} = 0 \quad \text{a.s. } \forall j, \tag{15}$$

which is the standard martingale difference noise assumption; see Kushner and Clark (1978, Example 1 in Section 2.2). Here, E denotes the expectation and a.s. means almost surely. The condition (15) is trivially satisfied when the noise is a sequence of zero-mean random vectors independent of θ_j's as considered in Blum (1954) and Kiefer and Wolfowitz (1952). An estimate of the gradient in (13) used by FDSA is thus given as

$$g_j(\theta_j) = \frac{1}{2c_j}\begin{bmatrix} y_j^{1+} - y_j^{1-} \\ \vdots \\ y_j^{m+} - y_j^{m-} \end{bmatrix}. \tag{16}$$

Assumption 9. The algorithm satisfies:

(i) $a_j, c_j > 0\ \forall k$; $a_j \to 0$, $c_j \to 0$ as $k \to \infty$, $\sum_{k=0}^{\infty} a_j = \infty$, $\sum_{k=0}^{\infty}(\frac{a_j}{c_j})^2 < \infty$; see (13), (14) and (16).
(ii) $\sup_j \|\theta_j\| < \infty$ a.s.
(iii) θ^* is an asymptotically stable solution of the differential equation

$$\frac{dz(t)}{dt} = -g(z). \tag{17}$$

(iv) Let $D(\theta^*) := \{z_0 \in \mathbb{R}^m : \lim_{t\to\infty} z(t, z_0) = \theta^*\}$, where $z(t, z_0)$ denotes the solution to the differential equation (17) with respect to the initial condition z_0, i.e. $D(\theta^*)$ is the domain of attraction. There exists a compact $S \subset D(\theta^*)$ such that $\theta_j \in S$ infinitely often for almost all sample points.

Remark 10 (*Spall, 1992, Section III.B*)**.** Contains remarks/justifications and references for the regularity conditions in Assumption 9. In particular, it is argued that Assumption 9(ii) is not restrictive in most applications.

Proposition 11. *Suppose Assumption 9 holds, then the output of the FDSA, $\theta_j \to \theta^*$ with probability one (w.p.1).*

Proof. The conditions for convergence stated here are taken from Kushner and Clark (1978, Thm. 2.3.5), which extend the original version in Blum (1954); see also Chin (1997). □

The remaining part of the subsection adapts the FDSA to carrying out the task of extremum seeking control of dynamical systems with output measurements corrupted by additive noise as illustrated in Fig. 4. Semi-global convergence to an extremum w.p.1 is established.

Consider the feedback configuration in Fig. 4 of a dynamical plant in Definition 1 satisfying Assumption 2 with steady-state input–output map $Q : \mathbb{R}^m \to \mathbb{R}$ and an extremum seeking algorithm taken to be the FDSA. These are interconnected through a T-periodic sampler (3) and a synchronised zero-order hold (2). In particular, for $j = 0, 1, \ldots$, during the jth FDSA algorithmic iteration, the output $\{u_k\}_{k=2mj}^{2m(j+1)-1}$ of the extremum seeker is defined in the following manner:

$$\begin{aligned} u_{2mj+2(i-1)} &= \theta_j + c_j e_i, \quad i = 1, \ldots, m \\ u_{2mj+2i-1} &= \theta_j - c_j e_i, \quad i = 1, \ldots, m. \end{aligned}$$

These values are held constant for T seconds and used as consecutive inputs to the dynamic plant. Recall the notation that the T-periodically sampled output of the plant corresponding to $\{u_k\}_{k=2mj}^{2m(j+1)-1}$ is $\{y_k\}_{k=2mj+1}^{2m(j+1)}$. Denote the noise-corrupted inputs to the FDSA by

$$\begin{aligned} y_j^{i+} &:= y_{2mj+2i-1} + \epsilon_{2mj+2i-1} \\ y_j^{i-} &:= y_{2mj+2i} + \epsilon_{2mj+2i} \end{aligned} \tag{18}$$

for $i = 1, \ldots, m$. The estimate of the derivative at θ_j can thus be made in accordance with (16) and the update θ_{j+1} (13). The following result is in order.

Theorem 12. *Suppose Assumption 9 holds and*

$$E\left\{\epsilon_{2mj+2i-1} - \epsilon_{2mj+2i} \middle| \theta_n;\ n \le j\right\} = 0 \quad a.s.$$
$$\forall i = 1, \ldots, m,\ j = 0, 1, \ldots.$$

Then given any $\Delta > 0$, there exists a sampling period $T > 0$ such that for any initial state $\|x_0\|_{\mathcal{A}(u_0)} \le \Delta$, $\theta_j \to \theta^*$ w.p.1, where $\theta^* \in \mathbb{R}^m$ satisfies $\nabla Q(\theta^*) = 0$.

Proof. Let $\nu > 0$ be such that $\nu < \frac{1}{2\sqrt{m}}$. By Lemma 6, there exists a sampling $T > 0$ such that

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

for any $\{u_k\}_{k=0}^{\infty} \subset \mathbb{R}^m$ and $\|x_0\|_{\mathcal{A}(u_0)} \le \Delta$. It follows that (18) can be written as

$$\begin{aligned} y_j^{i+} &:= Q(u_{2mj+2(i-1)}) + w_{2mj+2i-1} + \epsilon_{2mj+2i-1} \\ y_j^{i-} &:= Q(u_{2mj+2i-1}) + w_{2mj+2i} + \epsilon_{2mj+2i}, \end{aligned}$$

where $\{w_k\}_{k=1}^{\infty}$ is a real-valued sequence satisfying $|w_k| \le \nu$ for all $k = 1, 2, \ldots$. By (16), the FDSA update Eq. (13) can be rewritten as

$$\theta_{j+1} = \theta_j - a_j g(\theta_j) + a_j \Gamma_j + \frac{a_j}{2c_j} W_j + a_j \Upsilon_j, \tag{19}$$

where

$$\Gamma_j := g(\theta_j) - \frac{1}{2c_j}\begin{bmatrix} Q(u_{2mj}) - Q(u_{2mj+1}) \\ \vdots \\ Q(u_{2mj+2(m-1)}) - Q(u_{2mj+2m-1}) \end{bmatrix}$$

denotes the error which arises from using finite-difference estimation of the derivatives,

$$W_j := \begin{bmatrix} w_{2mj+1} - w_{2mj+2} \\ \vdots \\ w_{2m(j+1)-1} - w_{2m(j+1)} \end{bmatrix}$$

represents the perturbations on the steady-state input–output map by the dynamics of the plant, and

$$\Upsilon_j := \frac{1}{2c_j}\begin{bmatrix} \epsilon_{2mj+1} - \epsilon_{2mj+2} \\ \vdots \\ \epsilon_{2m(j+1)-1} - \epsilon_{2m(j+1)} \end{bmatrix}$$

represents the additive measurement noise term. Note that $\Gamma_j \to 0$ because $c_j \to 0$ as $j \to \infty$. Since $|w_k| \le \nu\ \forall k$,

$$\|W_j\|_2 \le 2\nu\sqrt{m} < 1 \quad \forall j. \tag{20}$$

Recall from Assumption 9(i) that $\sum_{k=0}^{\infty}(\frac{a_j}{c_j})^2 < \infty$. As such, this implies

$$\frac{a_j}{2c_j} W_j \to 0 \quad \text{as } j \to \infty. \tag{21}$$

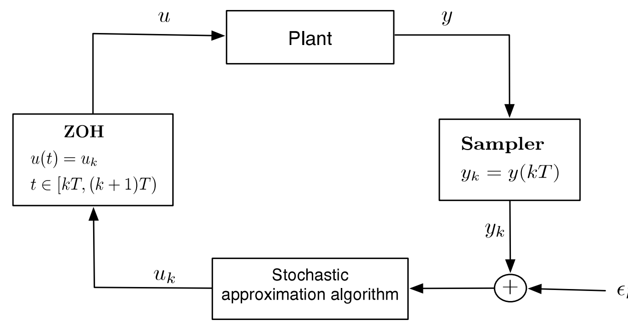


Fig. 4. Sampled-data extremum seeking control.

For $n = 0, 1, \ldots$, define $t_n := \sum_{j=0}^{n-1} a_j$,

$$W^0(t_n) := \sum_{j=0}^{n-1} \frac{a_j}{2c_j} W_j$$

$$W^0(t) := \frac{t_{n+1} - t}{a_n} W^0(t_n) + \frac{t - t_n}{a_n} W^0(t_{n+1}),$$

for $t \in (t_n, t_{n+1})$. Here W^0 is an interpolated function of the sequence $\{W_j\}_{j=0}^{\infty}$; see Kushner and Clark (1978, Section 2.1). Define also the sequence of left shifts:

$$W^n(t) := \begin{cases} W^0(t + t_n) - W^0(t_n) & t \ge -t_n \\ -W^0(t_n) & t < -t_n. \end{cases}$$

By (20) and (21), it follows that $\{W^n(\cdot)\}_{n=0}^{\infty}$ are uniformly continuous on $(-\infty, \infty)$, bounded on finite intervals in $(-\infty, \infty)$, and tend to zero as $n \to \infty$.

Now note that the FDSA update Eq. (19) for extremum seeking control of a dynamic plant differs from the standard Kiefer–Wolfowitz–Blum procedure (Kushner & Clark, 1978, Eq. (2.3.10)) only by the extra term $\frac{a_j}{2c_j}W_j$. By properties of $\{W^n(\cdot)\}_{n=0}^{\infty}$ just established, i.e. it does not affect the asymptotic behaviour of $\{\theta_j\}_{j=0}^{\infty}$, it follows that the proof methods of Kushner and Clark (1978, Thm. 2.3.1 and Thm. 2.3.5) are applicable here and yield the conclusion of the theorem.

In particular, Kushner and Clark (1978, Thm. 2.3.1) considers a sequence of update equations involving the left-shifted interpolated functions of the terms in (19), i.e. θ^n, Γ^n, and Υ^n, whose definitions reminisce that of W^n above; see Kushner and Clark (1978, Eq. (2.3.10)). The property in Assumption 9(i) and Kushner and Clark (1978, Lem. 2.2.1) applied to the assumption on the noise result in $\{\Gamma^n(\cdot)\}_{n=0}^{\infty}$ and $\{\Upsilon^n(\cdot)\}_{n=0}^{\infty}$ being uniformly continuous on $(-\infty, \infty)$, bounded on finite intervals in $(-\infty, \infty)$, and tending to zero as $n \to \infty$. As explained before, the extremum-seeking setting of this theorem introduces to the update equation an extra additive term W^n. By the property of $\{W^n(\cdot)\}_{n=0}^{\infty}$ established above, it follows that the proof of Kushner and Clark (1978, Thm. 2.3.1), and hence Kushner and Clark (1978, Thm. 2.3.5) holds here, whereby $\theta_j \to \theta^*$ w.p.1. □

Notice that the FDSA requires $2m$ measurements for every iteration of the algorithm. For systems with a large number of inputs, this can thus be very computationally costly and lead to slower convergence speed. This fact has motivated the developments of the RDSA and SPSA as alternatives which use fewer measurements. More specifically, only a pair of measurements are used per iteration of the algorithms.

5.2. Random directions (RDSA)

Let $\{d_j\}_{j=0}^{\infty}$ be a sequence of independent random vectors, each distributed uniformly over the surface of the m-dimensional sphere of radius m (Chin, 1997) (other probability distributions may also be used). The gradient in (13) is approximated by the RDSA as

$$g_j(\theta_j) = \frac{1}{2c_j} d_j(y_j^+ - y_j^-), \tag{22}$$

where $y_j^\pm = Q(\theta_j \pm c_j d_j) + \epsilon_j^\pm$, and the measurement noise satisfies

$$E\left\{\epsilon_j^+ - \epsilon_j^- \,\middle|\, \theta_i;\ d_i;\ \epsilon_{i-1}^\pm;\ i \le j\right\} = 0 \quad \text{a.s. and}$$

$$E[(\epsilon_j^\pm)^2] \le \alpha_0 \quad \text{a.s. } \forall j,$$

for some $\alpha_0 > 0$.

Assumption 13. $E(d_j d_j^T) = mI$ and there exists $\alpha_1 > 0$ such that $E[(d_j^i Q(\theta_j \pm c_j d_j))^2] \le \alpha_1$, for all j and $i = 1, 2, \ldots, m$, where the superscript T denotes the matrix transpose and I the identity matrix.

Proposition 14. *Suppose Assumptions 9 and 13 hold, then the output of the RDSA, $\theta_j \to \theta^*$ w.p.1.*

Proof. See Chin (1997) and Kushner and Clark (1978, Thm. 2.3.6). □

Consider now the sampled-data extremum seeking setup with measurement noise of Fig. 4 as in the last subsection, but with the RDSA deployed as the extremum seeking controller. In the jth RDSA algorithmic iteration, the output $\{u_k\}_{k=2j}^{2j+1}$ of the extremum seeker is defined by

$$u_{2j} = \theta_j + c_j d_j \quad \text{and} \quad u_{2j+1} = \theta_j - c_j d_j.$$

These give rise to the sampled plant's output $\{y_k\}_{k=2j+1}^{2(j+1)}$. Denote the noise-corrupted inputs to the RDSA by

$$\begin{aligned} y_j^+ &:= y_{2j+1} + \epsilon_{2j+1} \\ y_j^- &:= y_{2(j+1)} + \epsilon_{2(j+1)}. \end{aligned} \tag{23}$$

The estimate of the derivative at θ_j can thus be made according to (22) and the update θ_{j+1} can be made following (13).

Theorem 15. *Suppose Assumptions 9 and 13 hold,*

$$E\left\{\epsilon_{2j+1} - \epsilon_{2(j+1)} \,\middle|\, \theta_i;\ d_i;\ \epsilon_{2i-1};\ \epsilon_{2i};\ i \le j\right\} = 0 \quad \text{a.s.}$$

and $E[(\epsilon_j)^2] \le \alpha_0$ a.s. for some $\alpha_0 > 0$ and all j. Then given any $\Delta > 0$, there exists a sampling period $T > 0$ such that for any $\|x_0\|_{\mathcal{A}(u_0)} \le \Delta$, $\theta_j \to \theta^*$ w.p.1, where $\theta^* \in \mathbb{R}^m$ satisfies $\nabla Q(\theta^*) = 0$.

Proof. The assertion can be established by following the same arguments in the proof of Theorem 12 and appealing to Kushner and Clark (1978, Thm. 2.3.6), which is based on Kushner and Clark (1978, Thm. 2.3.6), for the convergence of the standard RDSA. In particular, the update equation for θ_j,

$$\begin{aligned} \theta_{j+1} = \theta_j - \frac{a_j}{2c_j} d_j(Q(u_{2j}) - Q(u_{2j+1}) \\ + \epsilon_{2j+1} - \epsilon_{2(j+1)} + w_{2j+1} - w_{2(j+1)}), \end{aligned}$$

differs from the standard RDSA by an extra term

$$\frac{a_j}{2c_j} d_j(w_{2j+1} - w_{2(j+1)}),$$

which tends to 0 as $j \to \infty$ since d_j is uniformly bounded, where $\{w_k\}_{k=1}^\infty$ is a real-valued sequence satisfying $|w_k| \le \nu$ for some $0 < \nu < \frac{1}{2m}$ and all $k = 1, 2, \ldots$ as in Theorem 12. □

In general, the RDSA does not have superior efficiency performance to the FDSA because the number of iterations may increase enough to nullify the decrease in the number of measurements per iteration (Kushner & Clark, 1978; Spall, 1992). The SPSA, on the other hand, is the preferable algorithm for systems with multiple inputs on both theoretical and numerical basis (Chin, 1997; Spall, 2003).

5.3. Simultaneous perturbation (SPSA)

Let $\Delta_j \in \mathbb{R}^m$ be a vector of m mutually independent zero-mean random variables $[\Delta_j^1, \Delta_j^2, \ldots, \Delta_j^m]$ and $\{\Delta_j\}_{j=0}^\infty$ a mutually independent sequence with Δ_j independent of $\theta_0, \theta_1, \ldots, \theta_j$. Δ_j^i can be symmetrically Bernoulli distributed about zero for example, as taken in the numerical studies of (Spall, 1992).

In the SPSA, the gradient in the optimisation procedure (13) is approximated by

$$g_j(\theta_j) = \frac{1}{2c_j} \begin{bmatrix} \dfrac{y_j^+ - y_j^-}{\Delta_j^1} \\ \vdots \\ \dfrac{y_j^+ - y_j^-}{\Delta_j^m} \end{bmatrix}, \tag{24}$$

where $y_j^\pm = Q(\theta_j \pm c_j \Delta_j) + \epsilon_j^\pm$, and the measurement noise terms satisfy

$$E\left\{\epsilon_j^+ - \epsilon_j^- \,\middle|\, \Delta_j;\ \theta_i;\ i \le j\right\} = 0 \quad \text{a.s. } \forall j$$

and for some $\alpha_0 > 0$, $E[(\epsilon_j^\pm)^2] \le \alpha_0$ a.s. for all j.

Assumption 16. There exist $\alpha_1, \alpha_2 > 0$ such that $E(Q(\theta_j \pm c_j \Delta_j)^2) \le \alpha_1$ and $E((\Delta_j^i)^{-2}) \le \alpha_2$ for all k and $i = 1, 2, \ldots, m$.

Proposition 17 (*Spall, 1992*)**.** *Suppose Assumptions 9 and 16 hold, then the output of the SPSA, $\theta_j \to \theta^*$ w.p.1.*

Consider now the sampled-data extremum seeking setup with measurement noise illustrated by Fig. 4 as in the previous subsections, but with the SPSA being the extremum seeking controller. In the jth SPSA algorithmic iteration, the output $\{u_k\}_{k=2j}^{2j+1}$ of the extremum seeker is defined by

$$u_{2j} = \theta_j + c_j \Delta_j \quad \text{and} \quad u_{2j+1} = \theta_j - c_j \Delta_j.$$

The corresponding noise-corrupted inputs to the SPSA are given again by (23). The estimate of the derivative at θ_j can thus be made according to (24) and the update θ_{j+1} can be computed following (13).

Theorem 18. *Suppose Assumptions 9 and 13 hold,*

$$E\left\{\epsilon_{2j+1} - \epsilon_{2(j+1)} \,\middle|\, \Delta_j;\ \theta_i;\ i \le j\right\} = 0 \quad \text{a.s. } \forall j$$

there exists an $\alpha_0 > 0$ such that $E[(\epsilon_j)^2] \le \alpha_0$ a.s. for all j, and $(\Delta_j^i)^{-1}$ is uniformly bounded for $i = 1, \ldots, m$ and sufficiently large j. Then given any $\Delta > 0$, there exists a sampling period $T > 0$ such that for any $\|x_0\|_{\mathcal{A}(u_0)} \le \Delta$, $\theta_j \to \theta^*$ w.p.1, where $\theta^* \in \mathbb{R}^m$ satisfies $\nabla Q(\theta^*) = 0$.

Proof. The claim can be shown by using the same arguments in the proof of Theorem 12 and appealing to Spall (1992, Prop. 1), which is based on Kushner and Clark (1978, Thm. 2.3.1), for the convergence of the standard SPSA. In particular, the update equation for θ_j,

$$\begin{aligned} \theta_{j+1} = \theta_j - \frac{a_j}{2c_j}(Q(u_{2j}) - Q(u_{2j+1}) \\ + \epsilon_{2j+1} - \epsilon_{2(j+1)} + w_{2j+1} - w_{2(j+1)}) \begin{bmatrix} \dfrac{1}{\Delta_j^1} \\ \vdots \\ \dfrac{1}{\Delta_j^m} \end{bmatrix}, \end{aligned}$$

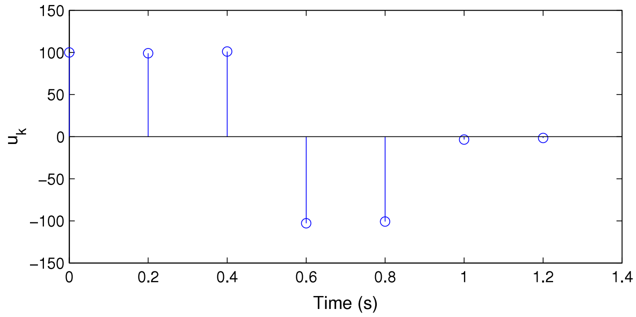


Fig. 5. FDSA output sequence (plant's input) for $T = 0.2$ s; $a_k = 1/k$.

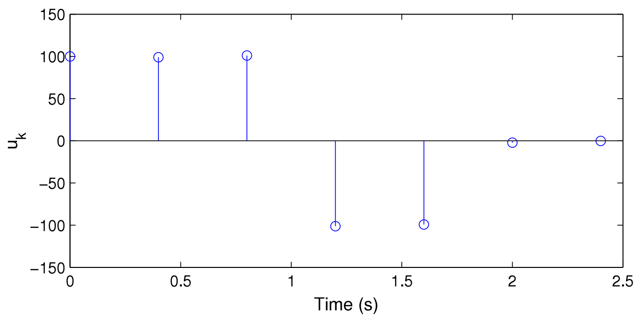


Fig. 6. FDSA output sequence for $T = 0.4$ s; $a_k = 1/k$.

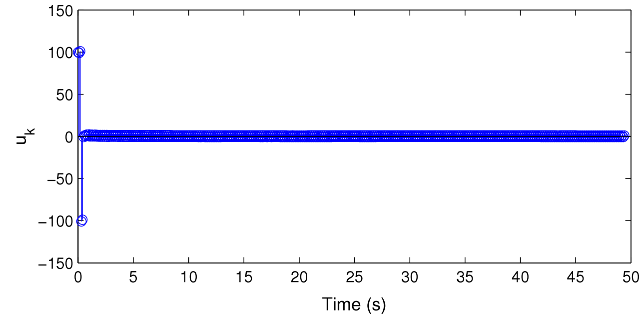


Fig. 7. FDSA output sequence for $T = 0.1$ s; $a_k = 1/k$.

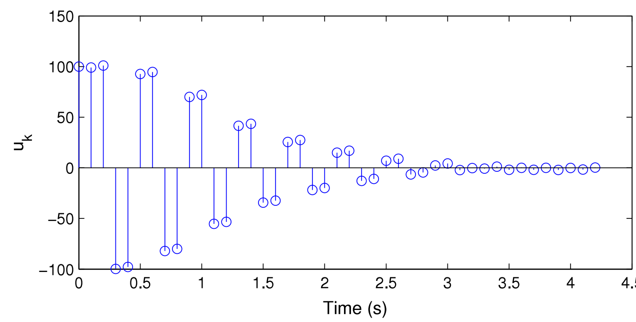


Fig. 8. FDSA output sequence for $T = 0.1$ s; $a_k = 1/k^{0.6}$.

differs from the standard SPSA by an additional term

$$\frac{a_j}{2c_j}(w_{2j+1} - w_{2(j+1)})\begin{bmatrix} \dfrac{1}{\Delta_j^1} \\ \vdots \\ \dfrac{1}{\Delta_j^m} \end{bmatrix},$$

which approaches 0 a.s. as $j \to \infty$, where $\{w_k\}_{k=1}^{\infty}$ is a real-valued sequence satisfying $|w_k| \le \nu$ for some $0 < \nu < \frac{1}{2\sqrt{m}}$ and all $k = 1, 2, \ldots$, similarly to the case in Theorem 12. □

Remark 19. In Spall (1992, 2003), it is shown that averaging over several gradient approximations (24) using conditionally (on θ_j) independent simultaneous perturbations at each iteration may provide benefits in terms of accuracy. Specifically, the gradient in (13) can be replaced by

$$g_j(\theta_j) = \frac{1}{q}\sum_{i=1}^{q} g_j^{(i)}(\theta_j), \quad q \ge 1,$$

where each $g_j^{(i)}(\theta_j)$ is generated as in (24) based on a different pair of measurements with simultaneous perturbations $\Delta_j^{(i)}$, which are independent conditionally on θ_j. The same idea can also benefit the RDSA; see Chin (1997).

6. Simulation examples

Consider the following one-dimensional nonlinear system with a single input:

$$\dot{x} = -x^3 + u^2, \quad x(0) = 2; \qquad y = x^3. \tag{25}$$

Note that for any fixed $u \in \mathbb{R}$, $x = u^{\frac{2}{3}}$ is a globally asymptotically stable equilibrium; see Section 2. It is apparent that the steady-state input–output map is $Q(u) = u^2$ with $u \in \Omega := (-200, 200)$, of which its unique global minimum is 0. The input is started at $u = 100$ and the FDSA is employed for minimum-seeking of the plant using the sampled-data control law detailed in Section 3. The sampled output measurements are corrupted by an i.i.d. zero-mean unity-variance Gaussian noise.

The gradient approximation step gain c_k is taken to be $1/k^{0.01}$, while the iteration step gain a_k is $1/k$. Note that the combinations of a_k and c_k satisfy Assumption 9 required for convergence. For the purpose of simulation, the algorithm is terminated when an input of magnitude less than 0.2 is found. Using sampling period of $T = 0.2$ s, it takes 1.2 s to locate an input estimate $u = 0.073$; see Fig. 5. By increasing the sampling period to $T = 0.4$ s, 2.4 s is needed to find the input $u = 0.0809$; see Fig. 6. If $T = 0.1$ s is used instead, 49.4 s is required to locate the input $u = 0.197$ (see Fig. 7).

For the case $T = 0.1$ s, by adjusting a_k to be $1/k^{0.6}$ and keeping c_k the same, the FDSA takes 4.2 s to locate the input $u = -0.1828$ (cf. Fig. 8). This shows that apart from the waiting time/sampling period T, the parameters in the stochastic algorithm itself also affects the accuracy/speed of convergence.

Now consider the following two-dimensional dynamical system with two inputs which is of the differential form in (1):

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

where

$$h(x_1, x_2) := 4x_1^2 + (x_2^3 - x_1)^2.$$

It follows that for any $u \in \mathbb{R}^m$, $x_1 = 0.5u_1$ and $x_2 = (0.5u_1 + u_2)^{\frac{1}{3}}$ is a globally asymptotically stable equilibrium. Thus, the steady-state map (cf. Definition 5) of the above dynamical system is given by the quadratic function

$$Q(u) = u_1^2 + u_2^2, \quad u \in \Omega := (-100, 100) \times (-100, 100),$$

which has a unique global minimum at $(0, 0)^T$. The sampled output measurements are corrupted by an i.i.d. zero-mean unity-variance Gaussian noise. Table 1 shows the number of samples required to locate an input of Euclidean norm less than 0.2 for FDSA, RDSA,

Table 1
Performance characteristics.

Algorithm	No. of samples taken	Estimated min
FDSA	244	$(0.044, -1.403)^T$
RDSA	220	$(-0.168, 0.055)^T$
SPSA	136	$(-0.099, -0.099)^T$

and SPSA algorithms initialised at $(50, 50)^T$. The sampling period is selected to be $T = 4$ s, and $c_k = 0.2/k^{0.01}$ and $a_k = 0.2/k^{0.6}$.

It can be seen from Table 1 that SPSA outperforms RDSA and FDSA for this example in terms of efficiency, as is consistent with the remark at the end of Section 5.2.

7. Conclusions

This paper applies stochastic optimisation methods to extremum seeking control of possibly infinite-dimensional time-invariant nonlinear systems with noisy output measurements and establishes semi-global convergence results. These contrast the standard gradient descent method based extremum seeking control scheme under noise-free measurements, where semi-global practical asymptotic stability with respect to an ultimate bound can be shown. Future research directions involve investigating generalisation of the stochastic approximation methods to multi-player and multi-agent settings. Different classes of stochastic optimisation methods may also be examined.

References

Ariyur, K. B., & Krstić, M. (2003). *Real-time optimization by extremum seeking control*. Wiley–Interscience.

Azuma, S., Sakar, M. S., & Pappas, G. (2012). Stochastic source seeking by mobile robots. *IEEE Transactions on Automatic Control*, *57*(9), 2308–2321.

Bertsekas, D. P., & Tsitsiklis, J. N. (1996). *Neuro-dynamic programming*. Athena Scientific.

Blum, J. (1954). Multidimensional stochastic approximation methods. *The Annals of Mathematical Statistics*, *25*, 737–744.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

Chin, D. C. (1997). Comparative study of stochastic algorithms for system optimization based on gradient approximations. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, *27*(2), 244–249.

Cruz-Hernández, C., Alvarez-Gallegos, J., & 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.

Khalil, H. K. (2002). *Nonlinear systems* (3rd ed.). Prentice Hall.

Khong, S. Z., Nešić, D., Manzie, C., & Tan, Y. (2013). Multidimensional global extremum seeking via the DIRECT optimisation algorithms. *Automatica*, *49*(7), 1970–1978.

Khong, S. Z., Nešić, D., Tan, Y., & Manzie, C. (2013). Unified frameworks for sampled-data extremum seeking control: global optimisation and multi-unit systems. *Automatica*, *49*(9), 2720–2733.

Kiefer, J., & Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function. *The Annals of Mathematical Statistics*, *23*, 462–466.

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

Kushner, H. J., & Clark, D. S. (1978). *Stochastic approximation methods for constrained and unconstrained systems*. Springer-Verlag.

Kushner, H. J., & Yin, G. G. (2003). *Stochastic approximation and recursive algorithms and applications*. Springer.

Liu, S.-J., & Krstić, M. (2012). *Stochastic averaging and stochastic extremum seeking*. Springer.

Manzie, C., & Krstić, M. (2009). Extremum seeking with stochastic perturbations. *IEEE Transactions on Automatic Control*, *54*(3), 580–585.

Nešić, D., Nguyen, T., Tan, Y., & Manzie, C. (2013). A non-gradient approach to global extremum seeking: an adaptation of the Shubert algorithm. *Automatica* *49*(3), 809–815.

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

Nusawardhana, , & Żak, S. H. (2004). Simultaneous perturbation extremum seeking method for dynamic optimization problems. In *Proceedings of American control conference* (pp. 2805–2810).

Polak, E. (1997). *Applied mathematical sciences*: *Vol. 124*. *Optimization—algorithms and consistent approximations*. Springer.

Popović, D., Janković, M., Magner, S., & Teel, A. R. (2006). Extremum seeking methods for optimization of variable cam timing engine operation. *IEEE Transactions on Control Systems Technology*, *14*(3), 398–407.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes: The art of scientific computing* (3rd ed.). Cambridge University Press.

Ruelle, D. (1989). *Elements of differentiable dynamics and bifurcation theory*. Academic Press.

Spall, J. C. (1992). Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, *37*(3), 332–341.

Spall, J. C. (2003). *Introduction to stochastic search and optimization*. Wiley–Interscience.

Stanković, M. S., & Stipanović, D. M. (2010). Extremum seeking under stochastic noise and applications to mobile sensors. *Automatica*, *46*(8), 1243–1251.

Strongin, R. G., & Sergeyev, Y. D. (2000). *Nonconvex optimization and its applications*, *Global optimization with non-convex constraints*. Kluwer Academic Publishers.

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

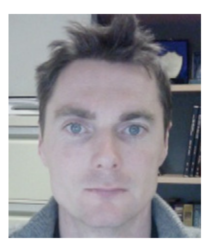
Teel, A. R., Moreau, L., & Nešić, D. (2003). A unification of time-scale methods for systems with disturbances. *IEEE Transactions on Automatic Control*, *48*, 1526–1544.

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

Zhang, C., & Ordóñez, R. (2011). *Advances in industrial control*, *Extremum-seeking control and applications: A numerical optimization-based approach*. Springer.

Sei Zhen Khong received the Bachelor of Electrical Engineering degree (with honours) and the Ph.D. degree from The University of Melbourne, Australia, in 2008 and 2012, respectively. He held a postdoctoral research fellowship in the Department of Electrical and Electronic Engineering, The University of Melbourne and is currently a postdoctoral researcher in the Department of Automatic Control, Lund University, Sweden. His research interests include distributed analysis of heterogeneous networks, robust control, linear systems theory, extremum seeking control, and sampled-data control.

Ying Tan received her Bachelor from Tianjin University, China in 1995. In 1998, she joined the National University of Singapore and obtained her Ph.D. in 2002. She joined McMaster University in 2002 as a postdoctoral fellow in the Department of Chemical Engineering. She has started her work in the Department of Electrical and Electronic Engineering, The University of Melbourne since 2004. Currently she is Future Fellow (2010–2013), which is a research position funded by the Australian Research Council. Her research interests are in intelligent systems, nonlinear control systems, real time optimisation, sampled-data distributed parameter systems and formation control.



Chris Manzie received the B.S. degree in Physics and the B.E. degree (with honours) in Electrical and Electronic Engineering and the Ph.D. degree from The University of Melbourne, Melbourne, Australia, in 1996 and 2001, respectively . Since 2003, he has been affiliated to the Department of Mechanical Engineering, The University of Melbourne, where he is currently an Associate Professor and an Australian Research Council Future Fellow. He was a Visiting Scholar with the University of California, San Diego in 2007, and a Visiteur Scientifique at IFP Energies Nouvelles, Paris in 2012. He has industry collaborations with companies including Ford Australia, BAE Systems, ANCA Motion and Virtual Sailing. His research interests lie in applications of model-based and extremum-seeking control in fields including mechatronics and energy systems. He is a member of the IEEE and IFAC Technical Committees on Automotive Control.

Dragan Nešić is a Professor in the Department of Electrical and Electronic Engineering (DEEE) at The University of Melbourne, Australia. He received his B.E. degree in Mechanical Engineering from The University of Belgrade, Yugoslavia in 1990, and his Ph.D. degree from Systems Engineering, RSISE, Australian National University, Canberra, Australia in 1997. Since February 1999 he has been with The University of Melbourne. His research interests include networked control systems, discrete-time, sampled-data and continuous-time nonlinear control systems, input-to-state stability, extremum seeking control, applications of symbolic computation in control theory, hybrid control systems, and so on. He was awarded a Humboldt Research Fellowship (2003) by the Alexander von Humboldt Foundation, an Australian Professorial Fellowship (2004–2009) and Future Fellowship (2010–2014) by the Australian Research Council. He is a Fellow of IEEE and a Fellow of IEAust. He is currently a Distinguished Lecturer of CSS, IEEE (2008-). He served as an Associate Editor for the journals Automatica, IEEE Transactions on Automatic Control, Systems and Control Letters and European Journal of Control.