2018 European Control Conference (ECC)
June 12-15, 2018. Limassol, Cyprus

Gradient-Based Extremum Seeking: Performance Tuning via Lie Bracket Approximations*

Christophe Labar[1], Jan Feiling[2] and Christian Ebenbauer[2]

Abstract— **In this paper, we propose model-free extremum seeking systems approximating a filtered-gradient descent law, instead of a simple gradient descent law. Namely, we consider that the gradient is low pass filtered before being fed in the gradient descent law. Exploiting the Lie bracket formalism, we derive general classes of systems that approximate the filtered-gradient descent law, and we focus on four particular schemes. The first ensures the boundedness of the update rates. The last three adapt the dither amplitude to enhance the steady state accuracy. The performances of those schemes are analyzed in simulation and compared with the performances of extremum seeking systems approximating a simple gradient descent law.**

I. INTRODUCTION

Model-free extremum seeking is a quite powerful class of real-time optimization methods. Typically, extremum seeking systems steer the input of a cost function towards a (local) optimizer, by combining time-periodic signals, the so-called *dithers*, with the on-line measurement of the cost.

The first paper on extremum seeking is probably the one authored by Leblanc in 1922 [1]. Since then, numerous approaches have been presented in the literature (see e.g. [2], [3] and [4] for a detailed survey). A first rigorous proof of local stability for nonlinear dynamical systems, based on averaging and singular perturbations, was presented in [5]. More recently, another way of analyzing model-free extremum seeking, based on Lie bracket approximation techniques, was proposed in [6]. This analysis highlighted the degree of freedom present on the choice of extremum seeking systems to approximate the gradient descent law. Such a degree of freedom has already been successfully exploited to improve the transient and steady state performances, e.g. the boundedness of the update rates presented in [7] or the asymptotic convergence stated in [8] and [9]. Furthermore, compared to the classical averaging technique, the Lie bracket formalism simplifies the design of such systems. It is for instance possible to analytically describe a class of systems that approximate the gradient descent law [9].

The existing extremum seeking schemes based on the Lie bracket formalism typically approximate a simple gradient descent law. However, it is known from the classical averaging analysis that the addition of a low pass filter in the closed-loop can improve both the transient performances and steady state accuracy. Therefore, the aim of this paper is to unify both approaches. Namely, we consider the case where the gradient is low pass filtered before being fed in the gradient descent law and we take benefit from the Lie bracket formalism to design systems approximating this filtered-gradient descent law. The benefit of introducing this low pass filter is twofold. First, by low pass filtering the estimated gradient, it attenuates the amplitude of the oscillations due to the presence of dithers. Then, it allows to design vector fields which are not admissible for approximating the simple gradient descent law.

A first contribution of the paper is to slightly generalize the first order Lie bracket approximation described in [6]. The main contribution is the design extremum seeking systems approximating the filtered-gradient descent law. General classes of systems are proposed and four schemes are studied more deeply. The first one ensures the boundedness of the update rates while the last three adapt the dither amplitude to enhance the steady state accuracy.

The rest of the paper is structured as follows. The notations and definitions are introduced in Section II. In Section III, we present the Lie bracket approximation. The problem addressed in this paper is formally stated in Section IV. Section V contains our main contribution.

II. NOTATIONS AND DEFINITIONS

$\mathbb{R}_{>0}$ and $\mathbb{Q}_{>0}$ represent the set of strictly positive real and rational numbers, respectively. $h'(x)$ stands for the first derivative of the scalar function $h : \mathbb{R} \to \mathbb{R}$. The Jacobian of a vector field $f : \mathbb{R}^n \to \mathbb{R}^m$ is denoted by $\frac{\partial f}{\partial x}$. The Lie bracket of two vector fields $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}^m$, denoted by $[f, g](x)$, is defined by $\frac{\partial g}{\partial x} f(x) - \frac{\partial f}{\partial x} g(x)$. Similarly to [10], we define semi-global practical uniform asymptotic stability of the origin as follows:

Definition 2.1: The origin of the n-dimensional system $\dot{x} = f(t, x, \alpha)$, with the vector of parameters $\alpha = [\alpha_1, \alpha_2, ..., \alpha_{n_\alpha}]^T$, is semi-Globally Practically Uniformly Asymptotically Stable (sGPUAS) if for every $r_B \in \mathbb{R}_{>0}$ and $r_V \in \mathbb{R}_{>0}$, there exist a $r_W \in \mathbb{R}_{>0}$, a $r_Q \in \mathbb{R}_{>0}$ and an $\alpha_1^* \in \mathbb{R}_{>0}$ such that for every $\alpha_1 \in (0, \alpha_1^*)$, there exists an $\alpha_2^* \in \mathbb{R}_{>0}$ such that for every $\alpha_2 \in (0, \alpha_2^*)$,..., there exist an $\alpha_{n_\alpha}^* \in \mathbb{R}_{>0}$ and a $t_1 \in \mathbb{R}$ such that for every $\alpha_{n_\alpha} \in (0, \alpha_{n_\alpha}^*)$ and $t_0 \in \mathbb{R}$ the following holds:

1) Boundedness: $||x(t_0)|| < r_V \Rightarrow ||x(t)|| < r_W$, $\forall t \geq t_0$;
2) Stability: $||x(t_0)|| < r_Q \Rightarrow ||x(t)|| < r_B$, $\forall t \geq t_0$;
3) Practical Convergence: $||x(t_0)|| < r_V \Rightarrow ||x(t)|| < r_B$, $\forall t \geq t_1 + t_0$.

*This work is supported by the Fonds National de la Recherche Scientifique under Grant ASP 24923120

[1]Christophe Labar is with the Department of Control Engineering and System Analysis, Université libre de Bruxelles, 1050 Brussels, Belgium chlabar@ulb.ac.be

[2]The authors are with the Institute for Systems Theory and Automatic Control, University of Stuttgart, 70569 Stuttgart, Germany {jan.feiling,ce}@ist.uni-stuttgart.de

Remark 2.1: If the vector of parameter α does not depend on r_V, sGPUAS reduces to global practical uniform asymptotic stability. Furthermore, if α does not depend on r_B, sGPUAS reduces to global uniform asymptotic stability.

III. THE LIE BRACKET APPROXIMATION

A first contribution of this work is a slight generalization of the first order Lie bracket approximation described in [6].

Consider the input-affine system

$$\dot{x}(t) = f_0(x(t)) + \sum_{i=1}^{l} \omega^{p_i} f_i(x(t)) u_i(k_i \omega t), \tag{1}$$

with $x \in \mathbb{R}^n$ the state vector, $u_i \in \mathbb{R}$ the control input, $\omega \in \mathbb{R}_{>0}$, $k_i \in \mathbb{Q}_{>0}$ and $p_i \in (0,1)$. In the sequel, we assume that the following holds:

Assumption 3.1: (Conditions on f_i and u_i)
For all $i \in \{0,...,l\}$ and $j \in \{1,...,l\}$:

1) $f_i(x) : \mathbb{R}^n \to \mathbb{R}^n$ is a class C^2 vector field, that may depend on a vector of parameters $\alpha \in \mathbb{R}^{n_\alpha}$;
2) $u_j(t) : \mathbb{R} \to \mathbb{R}$ is a bounded measurable function. For the sake of simplicity, and without loss of generality, it will be assumed that $\sup_{t\in\mathbb{R}} |u_j(t)| \leq 1$;
3) $u_j(t)$ is 2π-periodic, i.e. $u_j(t) = u_j(t + 2\pi)$, $\forall t \in \mathbb{R}$;
4) $u_j(t)$ has zero mean on one period, i.e. $\int_0^{2\pi} u_j(t)dt = 0$.

Assumption 3.2: Let T_{ij} be a common period to $u_i(k_i\omega t)$ and $u_j(k_j\omega t)$. If $p_i + p_j > 1$, then $[f_i, f_j](x) = 0, \forall x \in \mathbb{R}^n$ or $\int_0^{T_{ij}} \int_0^s u_j(k_j\omega s)u_i(k_i\omega p)\, dp\, ds = 0$. Furthermore, if $p_i + p_j + p_k \geq 2$, then $\frac{\partial}{\partial x}\left(\frac{\partial f_i(x)}{\partial x} f_j(x)\right) f_k(x) = 0, \forall x \in \mathbb{R}^n$.

Remark 3.1: Assumption 3.2 allows to select $p_i \in (0,1)$ and, hence, to extend the classical framework of first order Lie bracket approximations where $p_i = 0.5$ (see e.g. [6]). It will be made clearer in the sequel that such a degree of freedom allows to tune the performances of the extremum seeking systems.

Under Assumptions 3.1 and 3.2, a so-called *Lie bracket system* can be associated with (1), namely

$$\dot{\overline{x}} = f_0(\overline{x}) + \lim_{\omega\to\infty} \sum_{i=1}^{l-1} \sum_{j=i+1}^{l} [f_i, f_j](\overline{x})\gamma_{ij}(\omega), \tag{2}$$

with

$$\gamma_{ij}(\omega) := \frac{\omega^{p_i+p_j}}{T} \int_0^T \int_0^s u_j(k_j\omega s)u_i(k_i\omega p)\, dp\, ds \tag{3}$$

and T a common period of the control inputs, i.e. a common multiple of $\left(\frac{2\pi}{k_1\omega}, \frac{2\pi}{k_2\omega}, ..., \frac{2\pi}{k_l\omega}\right)$. Note that Assumptions 3.1 and 3.2 ensure that $\dot{\overline{x}}$ is finite, for all $\overline{x} \in \mathbb{R}^n$.

The stability properties of systems (1) and (2) are related by the following Lemma:

Lemma 3.1: Under Assumptions 3.1 and 3.2, if the origin is (semi-)globally (practically) uniformly asymptotically stable for system (2) with a vector of parameters $\alpha = [\alpha_1, \alpha_2, ..., \alpha_{n_\alpha}]^T$, then the origin is sGPUAS for system (1) with the vector of parameters $[\alpha^T, \omega^{-1}]^T$.

Proof: Due to the lack of space, the proof is not reported in this manuscript but can be found in [11]. ■

IV. PROBLEM STATEMENT

Consider a cost function $h(x) : \mathbb{R} \to \mathbb{R}$ satisfying the following assumption:

Assumption 4.1: i) $h(x)$ is a class C^2 function;
ii) $\exists x^* \in \mathbb{R} : h'(x)(x - x^*) > 0, \forall x \in \mathbb{R}\backslash\{x^*\}$.

Remark 4.1: Assumption 4.1 guarantees that x^* is the unique minimizer of $h(x)$ and that there is no other stationary point than x^*.

Remark 4.2: For the sake of clarity, single-input cost functions have been considered in this paper. Nevertheless, the proposed vector fields can be easily generalized to handle multi-input cost functions.

As mentioned in the introduction, most of the extremum seeking systems based on the Lie bracket formalism approximate a simple gradient descent law, namely

$$\dot{x} = -\rho h'(x), \tag{4}$$

with $\rho \in \mathbb{R}_{>0}$. However, the presence of dithers in the extremum seeking systems leads to oscillations on the estimated gradient. Accordingly, considering a gradient descent law in which the gradient is low pass filtered can be a first way to improve both the transient and steady state performances. The corresponding system is

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\rho y \\ -\omega_L(y - h'(x)) \end{bmatrix}, \tag{5}$$

with $\rho \in \mathbb{R}_{>0}$ and $\omega_L \in \mathbb{R}_{>0}$.

Remark 4.3: System (5) can also be interpreted as a particular case of the Heavy-ball method (see e.g. [12]).

The following Lemma states the stability properties of (5):

Lemma 4.1: Under Assumption 4.1, the point $(x^*, 0)$ is globally uniformly asymptotically stable for system (5).

Proof: The proof is reported in Appendix B of [11]. ■

Furthermore, the following result can be stated:

Lemma 4.2: The point $(x^*, 0)$ is sGPUAS, with parameter ω^{-1}, for any input-affine system (1) satisfying Assumptions 3.1 and 3.2 whose associated Lie bracket system is (5).

Proof: The proof follows from Lemmas 4.1 and 3.1. ■

The main aim of this paper is to propose vector fields and control inputs, consistent with Assumptions 3.1 and 3.2, such that system (1) approximates (5). The focus is set on vector fields allowing to improve, for a given dither frequency, the steady state accuracy. Note that Lemma 4.2 ensures that the point $(x^*, 0)$ is sGPUAS for the proposed schemes.

V. DESIGN OF EXTREMUM SEEKING SYSTEMS

In this section, we start by presenting the simplest input-affine system, in the form of (1), approximating (5) (Section V-A). Two slight modifications of this scheme are then proposed to improve the steady state accuracy (Sections V-B and V-C). In the spirit of [13], a system ensuring the boundedness of the update rates is presented in Section V-D. Three schemes with adaptive dither amplitudes are then proposed (Sections V-E - V-G). The section is concluded by describing general classes of input-affine systems whose associated Lie bracket system is (5) (Sections V-H and V-I).

In the proposed schemes, the inputs $u_1(t)$ and $u_2(t)$ are assumed to satisfy Assumption 3.1 and the constants $k_1 \in \mathbb{Q}_{>0}$ and $k_2 \in \mathbb{Q}_{>0}$ are such that $\gamma_{12}(\omega) \neq 0, \forall \omega \in \mathbb{R}_{>0}$. It can be shown (see [11]) that under those conditions, if $p_1 + p_2 = 1$ then $\gamma_{12}(\omega)$ is a non zero constant, $\forall \omega \in \mathbb{R}_{>0}$. In the sequel, this constant value is simply denoted by γ_{12}.

To compare the performances of the presented schemes, and to analyze the influence of their tuning parameters, the following case study is considered: $h(x) = (x-2)^2 + 100$, $x(t_0) = 5$, $y(t_0) = 0$, $\omega = 100\,\mathrm{rad/s}$, $k_1 = k_2 = 1$, $u_1(t) = \sin(t)$ and $u_2(t) = \cos(t)$ (hence $\gamma_{12} = -0.5$). Note that the results obtained with such a quadratic cost function also provide an insight into the results that would be obtained with an arbitrary function of class C^2, near the minimum.

A. The Basic Scheme

The simplest system, in the form of (1), approximating (5) is undoubtedly

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \begin{bmatrix}-\rho y\\ -\omega_L y\end{bmatrix} + \sqrt{\omega}\begin{bmatrix}1 & 0\\ 0 & \omega_L \frac{h(x)}{\gamma_{12}}\end{bmatrix}\begin{bmatrix}u_1(k_1\omega t)\\ u_2(k_2\omega t)\end{bmatrix}. \quad (6)$$

In Fig. 1, the trajectory of (6) is compared with the trajectory of the "simple" gradient-based extremum seeking scheme given in [6], namely $\dot{x} = -\sqrt{\omega}(\cos(\omega t) - 2\rho h(x)\sin(\omega t))$. It can be noticed that, as expected, the low pass filter enhances the performances. For a same convergence time, the amplitude of the oscillations are reduced on the entire trajectory and the steady state error is about 20 times smaller.

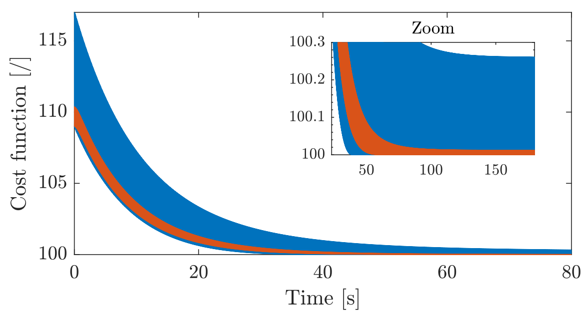

Fig. 1. Effect of adding a low pass filter: low pass filtered ES (6) (—); "simple" gradient descent ES (—); $\rho = 0.025\,\mathrm{rad/s}$, $\omega_L = 2\,\mathrm{rad/s}$

B. Balancing the Dither Amplitude

In this section, the vector fields introduced in (6) are slightly modified to further increase the steady state accuracy. The idea is to introduce a small parameter $\epsilon \in \mathbb{R}_{>0}$ to distribute the dithers amplitude between the x-dynamics and the input of the low pass filter. The resulting system is

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \begin{bmatrix}-\rho y\\ -\omega_L y\end{bmatrix} + \sqrt{\omega}\begin{bmatrix}\epsilon & 0\\ 0 & \epsilon^{-1}\omega_L \frac{h(x)}{\gamma_{12}}\end{bmatrix}\begin{bmatrix}u_1(k_1\omega t)\\ u_2(k_2\omega t)\end{bmatrix}. \quad (7)$$

Note that such a modification does not affect the associated Lie bracket system, which is still given by (5). From a practical point of view, (7) allows to tune the dither amplitude independently of the dither frequency. This provides an additional degree of freedom compared to the existing extremum seeking systems based on Lie bracket approximations. In Fig. 2, one can see that the amplitude of the oscillations can be reduced and the steady state accuracy can be improved by lowering the value of ϵ. However, as ϵ decreases, the trajectory of (7) follows less accurately the trajectory of (5). There is therefore a trade-off between the accuracy of approximation and the steady state accuracy. Furthermore, it is worth to mention that this tuning rule has some limitations: below a given value of ϵ, decreasing further ϵ leads to an increase of the oscillations and does not improve anymore the steady state accuracy. The smaller the value of ω_L and ρ for a given $h(x)$, the smaller the value of ϵ that can be used before an increase of the oscillations. In Fig. 3, it can be observed that, by reducing the value of ω_L, for the same values of ω and ϵ, the trajectory of (7) comes closer to the one of (5) and the steady state error is reduced. Note that this balancing of the dithers amplitude can be utilized each time the matrix $[f_1\ f_2]$ is triangular (see e.g. Section V-D).

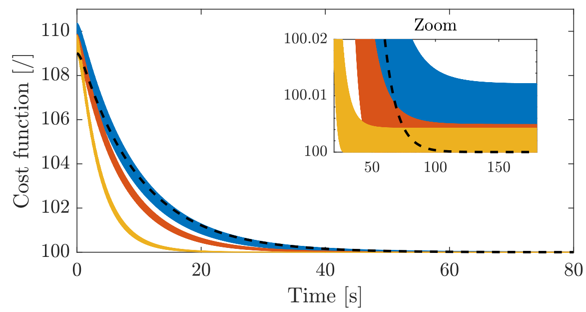

Fig. 2. Simulation of (7): $\epsilon = 1$ (—), $\epsilon = 0.5$ (—), $\epsilon = 0.25$ (—); Lie bracket system (5) (- -); $\rho = 0.025\,\mathrm{rad/s}$, $\omega_L = 2\,\mathrm{rad/s}$

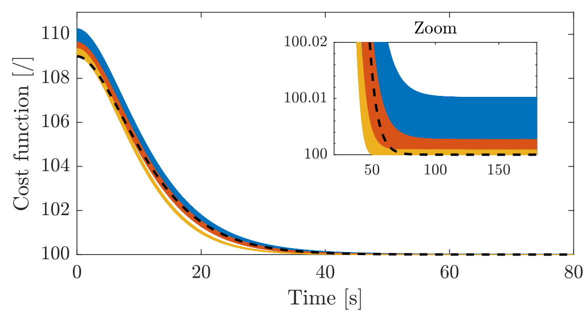

Fig. 3. Simulation of (7): $\epsilon = 1$ (—), $\epsilon = 0.5$ (—), $\epsilon = 0.25$ (—); Lie bracket system (5) (- -); $\rho = 0.025\,\mathrm{rad/s}$, $\omega_L = 0.2\,\mathrm{rad/s}$

C. Balancing the Dither Frequency

Referring to (2) and (3), one can conclude that if the vector fields and control inputs satisfy Assumptions 3.1 and 3.2, then selecting $p_i = p_j = 0.5$ or any $p_i, p_j \in (0,1)$ such that $p_i + p_j = 1$ leads to the same Lie bracket system. Such a degree of freedom on the choice of p_i, p_j can also be exploited to enhance the steady state accuracy. Let us study this strategy for the basic scheme (6), namely consider

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \begin{bmatrix}-\rho y\\ -\omega_L y\end{bmatrix} + \begin{bmatrix}1 & 0\\ 0 & \omega_L \frac{h(x)}{\gamma_{12}}\end{bmatrix}\begin{bmatrix}\omega^{p_1} u_1(k_1\omega t)\\ \omega^{1-p_1} u_2(k_2\omega t)\end{bmatrix}. \quad (8)$$

In Fig. 4, it can be seen that the results obtained with $p_1 \in (0, 0.5]$ provide a similar behavior to the one obtained by introducing the factor ϵ. This is not surprising since, for a given value of ω, both strategies are identical. However, from a stability point of view, the two tuning rules are different. In the previous case, ϵ was in the vector fields and had to be seen as a constant for every ω (cf. Lemma 3.1). In the present case, the "weighting factor" automatically adapts with ω, enhancing further the steady state accuracy. This strategy can be exploited each time the matrix $[f_1\ f_2]$ is triangular.

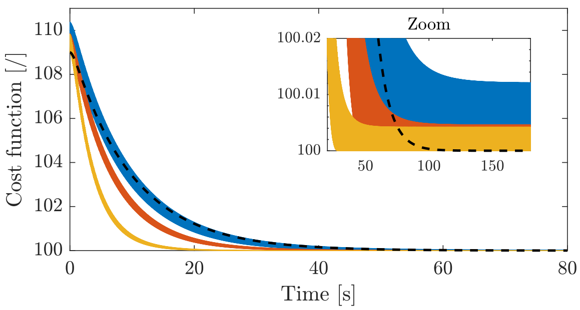

Fig. 4. Simulation of (8): $p = 1$ (—), $p = 1/3$ (—), $p = 1/5$ (—); Lie bracket system (5) (- -); $\rho = 0.025\,\mathrm{rad/s}$, $\omega_L = 2\,\mathrm{rad/s}$

D. Boundedness of the Update Rates

Motivated by [13], we propose the following system to ensure the boundedness of the update rates

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \begin{bmatrix}-\rho y\\ -\omega_L y\end{bmatrix} + \sqrt{\omega}\begin{bmatrix}\cos(h(x)) & 0\\ 0 & \frac{\omega_L}{\gamma_{12}}\sin(h(x))\end{bmatrix}\begin{bmatrix}u_1(k_1\omega t)\\ u_2(k_2\omega t)\end{bmatrix} + \sqrt{\omega}\begin{bmatrix}\sin(h(x)) & 0\\ 0 & -\frac{\omega_L}{\gamma_{34}}\cos(h(x))\end{bmatrix}\begin{bmatrix}u_3(k_3\omega t)\\ u_4(k_4\omega t)\end{bmatrix}, \quad (9)$$

where the $u_i(t)$ satisfy Assumption 3.1 and the $k_i \in \mathbb{R}_{>0}$ are such that i) $\gamma_{12}(\omega)$ and $\gamma_{34}(\omega)$ are non zero, $\forall\omega \in \mathbb{R}_{>0}$, ii) $\gamma_{ij}(\omega) = 0$, for $i, j \in \{1, ..., 4\}, (i, j) \notin \{(1, 2), (3, 4)\}$ and $\forall\omega \in \mathbb{R}_{>0}$. Note that the Lie bracket system associated with (9) is still (5). The boundedness of the update rates can be justified as follows. For any value of $h(x)$, the input of the low pass filter is bounded by $\sqrt{\omega}(\gamma_{12}^{-1} + \gamma_{34}^{-1})$. Provided $|y(t_0)| \leq \sqrt{\omega}(\gamma_{12}^{-1} + \gamma_{34}^{-1})$, it results $|\dot{y}(t)| \leq 2\omega_L\sqrt{\omega}(\gamma_{12}^{-1} + \gamma_{34}^{-1})$ and $|y(t)| \leq \sqrt{\omega}(\gamma_{12}^{-1} + \gamma_{34}^{-1})$, $\forall t \geq t_0$. Therefore, $|\dot{x}(t)| \leq \sqrt{\omega}(2 + \rho(\gamma_{12}^{-1} + \gamma_{34}^{-1}))$, $\forall t \geq t_0$. Note that, in this case, two first order brackets had to be combined to approximate (5), which is different from [13]. In Fig. 5, it can be seen that the result obtained with (9) (blue curve) does not improve the results obtained with the basic scheme (cf. Fig. 1). However, if we make use the weighting factor $\epsilon \in \mathbb{R}_{>0}$ introduced in Section V-B, namely if we consider

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \begin{bmatrix}-\rho y\\ -\omega_L y\end{bmatrix} + \sqrt{\omega}\begin{bmatrix}\epsilon\cos(h(x)) & 0\\ 0 & \epsilon^{-1}\frac{\omega_L}{\gamma_{12}}\sin(h(x))\end{bmatrix}\begin{bmatrix}u_1(k_1\omega t)\\ u_2(k_2\omega t)\end{bmatrix} + \sqrt{\omega}\begin{bmatrix}\epsilon\sin(h(x)) & 0\\ 0 & -\epsilon^{-1}\frac{\omega_L}{\gamma_{34}}\cos(h(x))\end{bmatrix}\begin{bmatrix}u_3(k_3\omega t)\\ u_4(k_4\omega t)\end{bmatrix}, \quad (10)$$

we can observe that the performances are improved (orange curve). The trajectory is now very close to the Lie bracket system and the steady state error is reduced by a factor 20. Furthermore, if we compare the trajectory of (10) with the trajectory of the bounded update rates scheme presented in [13] (red curve), we can see that the amplitude of the oscillations are reduced and the steady state error is 2.5 times smaller. Note that such an improvement would not have been possible without the presence of the low pass filter. Similarly to (7), for given values of ρ, ω_L and ω, there is a value of ϵ below which the performances deteriorate. Finally, it is worth to mention that similar performances can be obtained by using the tuning method described in Section V-C.

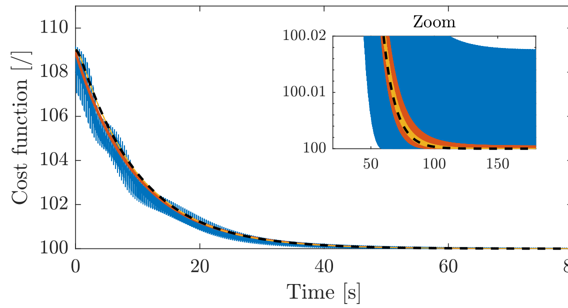

Fig. 5. Simulation of (10): $\epsilon = 1$ (—), $\epsilon = 0.05$ (—); Simulation of the "classical" bounded update rate scheme [13] (—); Lie bracket system (5) (- -); $\rho = 0.025\,\mathrm{rad/s}$, $\omega_L = 2\,\mathrm{rad/s}$, $u_1(t) = u_3(t) = \sin(t)$, $u_2(t) = u_4(t) = \cos(t)$, $k_1 = k_2 = 1$, $k_3 = k_4 = 1.2$

E. Zero Minimal Cost

When the minimal cost is zero [1], the following system can drastically improve the steady state performances.

$$\begin{bmatrix}\dot{x}\\ \dot{y}\end{bmatrix} = \begin{bmatrix}-\rho y\\ -\omega_L y\end{bmatrix} + \sqrt{\omega}\begin{bmatrix}\frac{\sqrt{h(x)}}{1+\sqrt{h(x)}} & 0\\ 0 & \omega_L\frac{h(x)+2\sqrt{h(x)}}{\gamma_{12}}\end{bmatrix}\begin{bmatrix}u_1(k_1\omega t)\\ u_2(k_2\omega t)\end{bmatrix}. \quad (11)$$

This system guarantees that the dither amplitude acting on the cost input is always strictly smaller than 1 and tends to zero as the cost approaches its minimum. The results depicted in Fig. 6, for $h(x) = (x - 2)^2$, suggest that (11) even ensures an asymptotic convergence to the minimum (the steady state error is smaller than 10^{-23} after 600s). A rigorous analysis of this result will be examined in a future work. One challenge of this analysis lies in the fact that the vector fields are not C^2. A first study of non-C^2 vector fields was done in [7]. However, the considered vector fields were non differentiable at a single point, while in the present case, they are not differentiable on the entire line (x^*, y), with $y \in \mathbb{R}$. In Fig. 6, we also compare trajectory of (11) with the trajectory of the scheme presented in [8], namely $\dot{x} = \sqrt{2\rho\omega h(x)}(\sin(\log(h(x)))\cos(\omega t) + \cos(\log(h(x)))\sin(\omega t))$. It can be observed that both curves are close to each other. Furthermore, one can see that (11) results in smaller maximal oscillations, up to 2.5 times smaller. Note that the parameter ϵ (cf. Section V-B) can also be introduced to further decrease the amplitude of those oscillations.

F. Adaptive Dither Amplitude Proportional to the Gradient

In this section, we modify the scheme presented in Section V-B so that ϵ decreases as $h'(x)$ approaches 0. This adaptation of ϵ allows to end up with much smaller values

[1] If the minimal cost is not zero but its value is known, or if a lower bound on the minimum is available, the proposed scheme can be used with $\tilde{h}(x) = h(x) - h_e^*$, where h_e^* denotes the minimum or the lower bound.

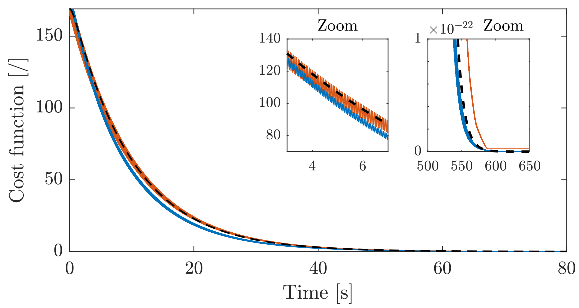

Fig. 6. Simulation of (11) (—), Simulation of [8] (—), Lie bracket system (5) (- -); $\rho = 0.025\,\mathrm{rad/s}$, $\omega_L = 2\,\mathrm{rad/s}$, $x(t_0) = 15$

of ϵ than the ones admissible for (7) and, hence, to enhance further the steady state accuracy. The resulting system is

$$\begin{bmatrix}\dot{x}\\ \dot{y}\\ \dot{v}\\ \dot{a}\end{bmatrix}=\begin{bmatrix}-\rho y\\ -\omega_L y\\ -\omega_H(v-h(x))\\ -c_a(a-|y|)\end{bmatrix}+\sqrt{\omega}\begin{bmatrix}\frac{\bar{\epsilon}+a}{1+a} & 0\\ 0 & \frac{\omega_L}{\gamma_{12}}(h(x)-v)\frac{1+a}{\bar{\epsilon}+a}\\ 0 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}u_1(k_1\omega t)\\ u_2(k_2\omega t)\end{bmatrix}, \tag{12}$$

with $\rho, \omega_L, \omega_H, c_a \in \mathbb{R}_{>0}$, $\bar{\epsilon} \in \mathbb{R}_{>0}$ an arbitrary small parameter and where the term $\frac{\bar{\epsilon}+a}{1+a}$ can be seen as $\epsilon(t)$. The following Lemma states the stability properties of (12):

Lemma 5.1: Under Assumption 4.1, the point $(x^*, 0, h(x^*), 0)$ is sGPUAS for system (12) with parameter ω^{-1}.

Proof: The proof is given in Appendix E of [11]. ■

The idea behind (12) can roughly be explained as follows. For ω sufficiently high, the input of the low pass filter is approximately the gradient (cf. the Lie bracket approximation). Therefore, for ρ sufficiently small, $y(t) \approx h'(x(t))$ and, hence, for c_a sufficiently high, $a(t) \approx |h'(x(t))|$. Accordingly, the dither amplitude, in the x-dynamics, tends to $\bar{\epsilon}$ as x approaches x^*. In Fig. 7, it can be observed that (12) allows to achieve steady state errors much smaller than the ones obtained with (7). One gets for instance an error of $6\,10^{-14}$ for $\bar{\epsilon} = 10^{-6}$. Note that, for all the cases, the trajectory follows accurately the trajectory of the Lie bracket system (black dashed curve). The simulations carried out suggest that, the smaller the selected $\bar{\epsilon}$, the smaller the c_a to obtain the best performances. Furthermore, for given values of ρ and ω_L, there seems to be a lower bound on the minimal error that can be achieved.

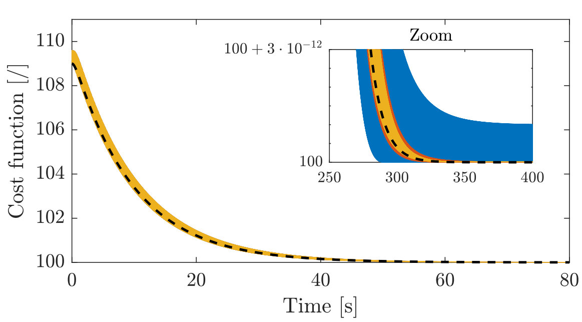

Fig. 7. Simulation of (12): $\bar{\epsilon} = 10^{-5}$ (—), $\bar{\epsilon} = 10^{-6}$ (—), $\bar{\epsilon} = 10^{-7}$ (—); Lie bracket system of (12)(- -), $\rho = 0.025\,\mathrm{rad/s}$, $\omega_L = 2\,\mathrm{rad/s}$

G. Adaptive Dither Amplitude With Saddle Point Dynamics

In this section, we adapt the adaptive dither amplitude scheme of [14] to the framework of this paper. Namely, we consider

$$\begin{bmatrix}\dot{x}\\ \dot{y}\\ \dot{z}\\ \dot{\nu}\end{bmatrix}=\begin{bmatrix}-\rho y+p(\nu)\sqrt{\omega}\epsilon(\cos(h(x))u_1(k_1\omega t)+\sin(h(x))u_3(k_3\omega t))\\ -\omega_L y+\omega_L\frac{p(\nu)\sqrt{\omega}}{\epsilon}\left(\sin(h(x))\frac{u_2(k_2\omega t)}{\gamma_{12}}-\cos(h(x))\frac{u_4(k_4\omega t)}{\gamma_{34}}\right)\\ a_1(\nu-1)\\ -a_2(\nu-1)+a_3(h(x)-z)\end{bmatrix}, \tag{13}$$

where the $u_i(t)$ satisfy Assumption 3.1 and the $k_i \in \mathbb{R}_{>0}$ are such that i) $\gamma_{12}(\omega)$ and $\gamma_{34}(\omega)$ are non zero, $\forall \omega \in \mathbb{R}_{>0}$, ii) $\gamma_{ij}(\omega) = 0$, for $i, j \in \{1, \ldots, 4\}, (i, j) \notin \{(1, 2), (3, 4)\}$ and $\forall \omega \in \mathbb{R}_{>0}$. By adapting the dither amplitude, the proposed scheme allows to end up with arbitrarily small dithers amplitude in a neighborhood of the minimum. The idea behind (13) is the following. First, we rewrite the minimization of $h(x)$ as the constrained minimization problem: $\min_{\{x,z\}\in\mathbb{R}^2} z$ such that $h(x) - z = 0$. We then associate the Lagragian $L(x, z, \nu) = z + \nu(h(x) - z)$, and consider the saddle point dynamics

$$\begin{bmatrix}\dot{x}\\ \dot{z}\\ \dot{\nu}\end{bmatrix} = \begin{bmatrix}-\nabla_x L(x, z, \nu)\\ -\nabla_z L(x, z, \nu)\\ \nabla_\nu L(x, z, \nu)\end{bmatrix} = \begin{bmatrix}-\nu h'(x)\\ -1 + \nu\\ h(x) - z\end{bmatrix}. \tag{14}$$

Since the equilibrium of (14) is $(x^*, h(x^*), 1)$, we know that ν always converges to 1. Therefore, we introduce the adaptive gain $p(\nu)$, a convex function with a minimum $p(1) = \bar{\epsilon} \in \mathbb{R}_{>0}$ that can be chosen arbitrary small. Additionally, we modify the ν-dynamics to ensure asymptotic convergence (see [14]) and introduce the constants $a_1, a_2, a_3 \in \mathbb{R}_{>0}$. Furthermore, following the motivation of this paper, we use a low-pass filter version of the gradient, resulting in the following dynamics

$$\begin{bmatrix}\dot{x}\\ \dot{y}\\ \dot{z}\\ \dot{\nu}\end{bmatrix} = \begin{bmatrix}-\rho y\\ -\omega_L(y - p^2(\nu)h'(x))\\ a_1(\nu - 1)\\ -a_2(\nu - 1) + a_3(h(x) - z)\end{bmatrix}, \tag{15}$$

which is nothing but the Lie bracket system of (13). In Fig. 8, it can be observed that the transient behavior of (13) is close to the one of the Lie bracket system, thanks to the choice of a small ϵ (curves are overlapped), like in Fig. 5.

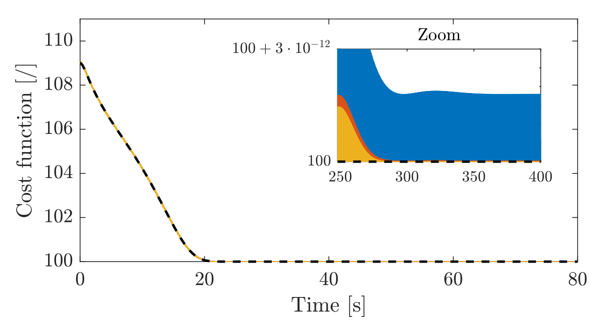

Fig. 8. Simulation of (13): $\bar{\epsilon} = 10^{-4}$ (—), $\bar{\epsilon} = 10^{-5}$ (—), $\bar{\epsilon} = 10^{-6}$ (—); Lie bracket system (15)(- -); $p(\nu) = \bar{\epsilon} + (\nu - 1)^2$, $\epsilon = 0.1$, $\rho = 0.025\,\mathrm{rad/s}$, $\omega_L = 2\,\mathrm{rad/s}$, $u_1(t) = u_3(t) = \sin(t)$, $u_2(t) = u_4(t) = \cos(t)$, $k_1 = k_2 = 1$, $k_3 = k_4 = 1.2$, $a_1 = 1/10$, $a_2 = 1/20$, $a_3 = 1/40$

Note also that the steady state error can be made arbitrarily small by decreasing $\bar{\epsilon}$. The role of a_1, a_2, a_3 is to tune the dynamics of ν and $p(\nu)$. Besides the specific tuning rules,

A		B ($\omega_L = 2rad/s$)		B ($\omega_L = 0.2rad/s$)		C		D		E	F		G	
-	SSE	ϵ	SSE	ϵ	SSE	p	SSE	ϵ	SSE	SSE	$\bar{\epsilon}$	SSE	$\bar{\epsilon}$	SSE
[6]	$2.2\,10^{-1}$	1	$4.3\,10^{-3}$	1	$3.6\,10^{-3}$	1	$8.0\,10^{-3}$	1	$1.1\,10^{-4}$	$<10^{-24}$	10^{-5}	$1.0\,10^{-12}$	10^{-4}	$1.8\,10^{-12}$
(6)	$8.0\,10^{-3}$	1/2	$1.7\,10^{-3}$	1/2	$9.6\,10^{-4}$	1/3	$3.0\,10^{-3}$	1/20	$1.4\,10^{-6}$		10^{-6}	$6.0\,10^{-14}$	10^{-5}	$1.8\,10^{-14}$
		1/4	$1.4\,10^{-3}$	1/4	$3.0\,10^{-4}$	1/5	$2.7\,10^{-3}$	[13]	$7.6\,10^{-5}$		10^{-7}	$3.8\,10^{-14}$	10^{-6}	$1.8\,10^{-16}$

TABLE I : Comparison of the maximal steady state error (SSE) for the different schemes.

one main result of [14] is the rigorous proof to achieve an arbitrary small steady state error of the non-filtered version of (13). The latter could be extended as follows: for every initial state $[x_0, y_0, z_0, \nu_0] \in \mathbb{R}^4$, there exist an $\omega^* > 0$ and $\bar{\epsilon}^* > 0$, such that for every $\omega > \omega^*$ and *every* $\bar{\epsilon} < \bar{\epsilon}^*$, the point $(x^*, 0, h(x^*), 1)$ is stable for system (13) and $\exists t_1 \in \mathbb{R}$: $|x(t) - x^*| < \bar{\epsilon}\sqrt{\omega}(2\epsilon + \frac{\rho}{\epsilon}(\gamma_{12}^{-1} + \gamma_{34}^{-1})), \forall t \geq t_1$. Notice that $\bar{\epsilon} > 0$ can be selected arbitrary small.

H. General Conditions on the Vector Fields

In the previous sections, we proposed input-affine systems to approximate (5). In the spirit of [9], the following Lemma generalizes those results, by giving general conditions on the vector fields and dither signals.

Lemma 5.2: Let $p_1 \in (0,1)$. Furthermore, let $u_1(t)$ and $u_2(t)$ satisfy Assumption 3.1 and $k_1 \in \mathbb{Q}_{>0}$ and $k_2 \in \mathbb{Q}_{>0}$ be such that $\gamma_{12}(\omega) \neq 0, \forall \omega \in \mathbb{R}_{>0}$. Let $g_1(x)$ and $g_2(x)$ be two functions of class C^2, with $g_1(x) \neq 0$, $\forall x \in \mathbb{R}$. Under Assumption 4.1, if $f_1(x)$ and $f_2(x)$ are such that

$$f_1(x) = \begin{bmatrix} 0 \\ -\omega_L \int \frac{\gamma_{12}^{-1}}{g_1(x)} + c \end{bmatrix} \text{ and } f_2(x) = \begin{bmatrix} g_1(x) \\ g_2(x) \end{bmatrix}$$

or

$$f_1(x) = \begin{bmatrix} g_1(x) \\ \omega_L g_2(x) \end{bmatrix} \text{ and } f_2(x) = \begin{bmatrix} k g_1(x) \\ \omega_L \left(k g_2(x) + c + \int \frac{\gamma_{12}^{-1}}{g_1(x)} \right) \end{bmatrix}$$

with $c \in \mathbb{R}$ and $k \in \mathbb{R}$, then (5) is the Lie bracket system associated with

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\rho y \\ -\omega_L y \end{bmatrix} + [f_1(h(x)) \ \ f_2(h(x))] \begin{bmatrix} \omega^{p_1} u_1(k_1 \omega t) \\ \omega^{1-p_1} u_2(k_2 \omega t) \end{bmatrix}. \quad (16)$$

Proof: The proof follows from direct computations. ■ Note that, under the assumptions of Lemma 5.2, Lemma 4.2 guarantees that the point $(x^*, 0)$ is sGPUAS for system (16).

I. General Form With Two First Order Lie Brackets

In this section, we extend what was done in Section V-D. Namely, we propose a way to adapt the vector fields designed for the simple gradient descent law to approximate (5).

Lemma 5.3: Let $\epsilon \in \mathbb{R}_{>0}$, $p_1 \in (0,1)$, $p_3 \in (0,1)$, $p_2 = 1 - p_1$ and $p_4 = 1 - p_3$. Furthermore, for $i \in \{1, ..., 4\}$, let $u_i(t)$ satisfy Assumption 3.1 and $k_i \in \mathbb{R}_{>0}$ be such that i) $\gamma_{12}(\omega)$ and $\gamma_{34}(\omega)$ are non zero, $\forall \omega \in \mathbb{R}_{>0}$, ii) $\gamma_{ij}(\omega) = 0$, for $i, j \in \{1, ..., 4\}, (i,j) \notin \{(1,2), (3,4)\}$ and $\forall \omega \in \mathbb{R}_{>0}$. Let $g_1(x), g_2(x) : \mathbb{R} \to \mathbb{R}$ be two functions of class C^2 such that $[g_1, g_2](h(x)) = h'(x)$. Under Assumption 4.1, if $f_1(x) := [\epsilon g_1(x) \quad 0]^T$, $f_2(x) := [0 \quad \epsilon^{-1} \gamma_{12}^{-1} \omega_L g_2(x)]^T$, $f_3(x) := [\epsilon g_2(x) \quad 0]^T$ and $f_4(x) := [0 \quad -\epsilon^{-1} \gamma_{34}^{-1} \omega_L g_1(x)]^T$, then (5) is the Lie bracket system associated with

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -\rho y \\ -\omega_L y \end{bmatrix} + \sum_{i=1}^{4} f_i(h(x)) \omega^{p_i} u_i(k_i \omega t). \quad (17)$$

Proof: The proof follows from direct computations. ■ Note that the parameter ϵ introduced in the vector fields is nothing but the tuning parameter presented in Section V-B. Note also that, under the assumptions of Lemma 5.3, Lemma 4.2 ensures that the point $(x^*, 0)$ is sGPUAS for system (17).

VI. CONCLUSION

The main aim of this paper was to study how the introduction of a low pass filter in a gradient descent law could be exploited to design new extremum seeking systems. General classes of systems were proposed and four schemes were studied more deeply, based on a case study. The obtained steady state errors are summarized in Table I. It can be concluded that the proposed schemes improve the results of the classical extremum seeking. Furthermore, by an adequate tuning of the parameters, they even allow to enhance the performances obtained by more complex schemes, like the ones presented in [8] and [13]. Future works will include the rigorous analysis of the stability properties of (11) and (13).

REFERENCES

[1] M. Leblanc, “Sur l’électrification des chemins de fer au moyen de courants alternatifs de fréquence élevée,” in *Revue Générale de l’ électricité*, 1922.

[2] K. T. Atta and M. Guay, “Fast proportional integral phasor extremum seeking control for a class of nonlinear system,” in *20th IFAC World Congress*, 2017.

[3] B. Hunnekens, M. Haring, N. van de Wouw, and H. Nijmeijer, “A dither-free extremum-seeking control approach using 1st-order least-squares fits for gradient estimation,” in *53rd CDC*, 2014.

[4] Y. Tan, W. Moase, C. Manzie, D. Nešić, and I. Mareels, “Extremum seeking from 1922 to 2010,” in *29th Chinese Control Conference (CCC), 2010*. IEEE, 2010, pp. 14–26.

[5] M. Krstić and H.-H. Wang, “Stability of extremum seeking feedback for general nonlinear dynamic systems,” *Automatica*, vol. 36, no. 4, pp. 595–601, 2000.

[6] H.-B. Dürr, M. S. Stanković, C. Ebenbauer, and K. H. Johansson, “Lie bracket approximation of extremum seeking systems,” *Automatica*, vol. 49, no. 6, pp. 1538 – 1552, 2013.

[7] A. Scheinker and M. Krstić, “Non-C^2 Lie bracket averaging for nonsmooth extremum seekers,” *Journal of Dynamic Systems, Measurement, and Control*, vol. 136, 2013.

[8] R. Suttner and S. Dashkovskiy, “Exponential stability for extremum seeking control systems,” in *20th IFAC World Congress*, 2017.

[9] V. Grushkovskaya, A. Zuyev, and C. Ebenbauer, “On a class of generating vector fields for the extremum seeking problem: Lie bracket approximation and stability properties,” *ArXiv*.

[10] S. P. Bhat and R. Cowlagi, “Semi-global practical stability of periodic time-varying systems via averaging: A lyapunov approach,” in *Proceedings of the 45th IEEE CDC*, 2006, pp. 361–365.

[11] C. Labar, J. Feiling, and C. Ebenbauer, “Gradient-based extremum seeking: Performance tuning via Lie bracket approximations,” *Internal Report IST, Stuttgart. Available at http://www.ist.uni-stuttgart.de/forschung/pdfs./GB_ES.pdf*, 2017.

[12] B. Polyak, *Introduction to Optimization*. Optimization Software, 1987.

[13] A. Scheinker and M. Krstić, “Extremum seeking with bounded update rates,” *Systems & Control Letters*, vol. 63, pp. 25 – 31, 2014.

[14] J. Feiling and C. Ebenbauer, “Extremum seeking with adaptive and arbitrary small steady state dither amplitude,” *under preparation*, 2017.