Saddle Point Seeking for Convex Optimization Problems ⋆

Hans-Bernd Dürr, Chen Zeng, Christian Ebenbauer

Institute for Systems Theory and Automatic Control, University of Stuttgart, Germany, (e-mail: {hans-bernd.duerr,chen.zeng,ce}@ist.uni-stuttgart.de)

Abstract: In this paper, we consider convex optimization problems with constraints. By combining the idea of a Lie bracket approximation for extremum seeking systems and saddle point algorithms, we propose a feedback which steers a single-integrator system to the set of saddle points of the Lagrangian associated to the convex optimization problem. We prove practical uniform asymptotic stability of the set of saddle points for the extremum seeking system for strictly convex as well as linear programs. Using a numerical example we illustrate how the approach can be used in distributed optimization problems.

1. INTRODUCTION

Extremum seeking is a control algorithm which allows to steer a system to the extremum of a function whose analytic expression is unknown. For many decades it has been used in various applications (see Tan et al. [2010], Krstić and Ariyur [2003], King et al. [2006], Guay et al. [2004]). Extremum seeking can be interpreted as gradient-free optimization algorithm for static maps (see Dürr et al. [2013]). There are only a few references dealing with extremum seeking for convex optimization problems with constraints (see DeHaan and Guay [2005], Frihauf et al. [2012], Poveda and Quijano [2012], Coito et al. [2005]). One can think of many applications where the goal is to minimize an unknown function and take constraints into account whose analytic expression are also unknown.

Consider for example a group of autonomous agents. Each of them shall minimize the distance between their position and a base station but at the same time they shall not exceed given distances among each other. This setup can be formulated such that the optimal positions are the solution to a convex optimization problem with constraints (see also Brunner et al. [2012]). While on the one hand there are many ways to measure the distance between two agents, it is on the other hand difficult to measure the gradient of the distance, especially when both agents are moving at the same time. Thus, for this problem setup one can try to find a control law where the gradient of the distances is not explicitly needed.

Generally speaking, we consider the class of problems where a convex function is minimized under convex constraints, i.e.

$$\begin{aligned} &\inf_x \; f(x) \\ &\text{s.t. } g_i(x) \leq 0, i = 1, \dots, m \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$, $f \in C^2 : \mathbb{R}^n \to \mathbb{R}$, $g_i \in C^2 : \mathbb{R}^n \to \mathbb{R}$, f, g_i convex. The goal is to propose a controller that steers a dynamical system to the solution of problem (1) without knowledge of the gradients of the functions f and g_i, $i = 1, \dots, m$.

⋆ This work was supported by the Deutsche Forschungsgemeinschaft (Emmy-Noether-Grant, Novel Ways in Control and Computation, EB 425/2-1, and Cluster of Excellence in Simulation Technology, EXC 310/1)

There are many different approaches in the literature dealing with algorithms for convex optimization problems (1). In this paper we focus especially on continuous-time saddle point algorithms (see Arrow et al. [1958], Dürr and Ebenbauer [2011], Nazemi [2012]). These algorithms exploit the fact that if (x^*, λ^*) is a saddle point of the Lagrangian

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) \tag{2}$$

associated to (1), then x^* is also a solution to the problem (1). A saddle point is defined as $L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*)$ for all $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m_+$. The idea of a saddle point algorithm is to minimize the Lagrangian with respect to x and to maximize it with respect to λ while assuring that λ stays nonnegative. A major advantage of saddle point algorithms is that they are well suited for distributed optimization problems (see Nedić and Ozdaglar [2009]). However, many of the proposed algorithms assume that the gradients of the functions f and g_i, $i = 1, \dots, m$ are available. Since we impose above that neither the gradients of the objective function f nor of the constraints g_i, $i = 1, \dots, m$ are known, these algorithms are not directly applicable in the scenario above. For this purpose, the proposed algorithm in this paper is a combination of extremum seeking (see e.g. Krstić and Ariyur [2003]) and saddle point algorithms (see e.g. Dürr and Ebenbauer [2011]).

We propose an extremum seeking feedback for single-integrator systems which are steered to the solutions of strictly convex as well as linear optimization problems. We impose, that the functions f and g_i, $i = 1, \dots, m$ in (1) are unknown as analytic expression. Instead, they can only be evaluated at a certain value of x, i.e. it is not possible to calculate the gradients as analytic expression. We establish practical stability of the overall systems and conclude the results with a numerical example where a distributed optimization setup is considered.

The remainder of this paper is structured as follows. In Section 2 we recall some mathematical preliminaries. In Section 3 we state our main result. We consider the case of (1) being strictly convex as well as (1) being a linear program. In Section 4 we consider a distributed optimization problem and show a numerical example. In Section 5 we summarize the results and state further research directions.

2. PRELIMINARIES

2.1 Notation

We will make use of the following notation. $\mathbb{Q}_{++}$ denotes the set of positive rational numbers. The intervals of real numbers are denoted by $(a,b) = \{x \in \mathbb{R} : a < x < b\}$, $[a,b) = \{x \in \mathbb{R} : a \leq x < b\}$ and $[a,b] = \{x \in \mathbb{R} : a \leq x \leq b\}$. The closure of a set $\mathcal{R} \subseteq \mathbb{R}^n$ is denoted by $\bar{\mathcal{R}}$. The norm $|\cdot|$ denotes the Euclidian norm. The Jacobian of a continuously differentiable function $b : \mathbb{R}^n \to \mathbb{R}^m$ is denoted by

$$\frac{\partial b(x)}{\partial x} := \begin{bmatrix} \frac{\partial b_1(x)}{\partial x_1} & \cdots & \frac{\partial b_1(x)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial b_m(x)}{\partial x_1} & \cdots & \frac{\partial b_m(x)}{\partial x_n} \end{bmatrix}$$

and the gradient of a continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is denoted by $\nabla_x f(x) := \left[\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial J(x)}{\partial x_n}\right]^\top$. The Lie bracket of two continuously differentiable vector fields $b_1, b_2 : \mathbb{R}^n \to \mathbb{R}^n$ is defined by $[b_1, b_2](x) := \frac{\partial b_2(x)}{\partial x} b_1(x) - \frac{\partial b_1(x)}{\partial x} b_2(x)$. We use $s \in \mathbb{C}$ for the complex variable of the Laplace transformation.

2.2 Extremum Seeking

In this section, we review some results from Dürr et al. [2013]. As it is shown in Dürr et al. [2013], certain extremum seeking systems can be written as input-affine systems

$$\dot{x} = b_0(x) + \sum_{i=1}^{m} b_i(x)\sqrt{\omega}u_i(\omega t) \tag{3}$$

with $x(t_0) = x_0 \in \mathbb{R}^n$ and $\omega > 0$.

We impose the following assumptions on b_i and u_i:

- A1 $b_i \in C^2 : \mathbb{R}^n \to \mathbb{R}^n$, $i = 0, ..., m$.
- A2 $u_i \in C^0 : \mathbb{R} \to \mathbb{R}$, $i = 1, ..., m$ and for every $i = 1, ..., m$ there exist constants $M_i > 0$ such that $\sup_{\theta \in \mathbb{R}} |u_i(\theta)| \leq M_i$.
- A3 $u_i(\cdot)$ is T-periodic, i.e., $u_i(\theta + T) = u_i(\theta)$ and has zero average, i.e., $\int_0^T u_i(\tau)d\tau = 0$, with $T > 0$ for all $\theta \in \mathbb{R}$, $i = 1, ..., m$.

These assumptions make sure that the results in Dürr et al. [2013] can be applied here. One main result of that paper is the approximation of trajectories of (3) by the trajectories of a so-called *Lie bracket system*

$$\dot{\bar{x}} = b_0(\bar{x}) + \sum_{\substack{i=1 \\ j=i+1}}^{m} [b_i(\bar{x}), b_j(\bar{x})]\nu_{ji}, \tag{4}$$

where

$$\nu_{ji} = \frac{1}{T}\int_0^T u_j(\theta)\int_0^\theta u_i(\tau)d\tau d\theta. \tag{5}$$

Let $\mathcal{I}, \mathcal{S} \subseteq \mathbb{R}^n$, $\mathcal{S}$ compact and $\bar{\mathcal{I}} \cap \mathcal{S}$ be non-empty. The following Lemma establishes semi-global practical asymptotic stability of (3) with respect to $\mathcal{I}$ assuming that (3) is uniformly positively invariant with respect to $\mathcal{I}$ and (4) is globally asymptotically stable with respect to $\mathcal{I}$. The stability definitions are recalled below.

Lemma 1. Let Assumptions A1 – A3 be satisfied and suppose that a compact set $\mathcal{S}$ is globally uniformly asymptotically stable with respect to $\mathcal{I}$ for (4). Suppose furthermore that the solutions of (3) are uniformly positively invariant with respect to $\mathcal{I}$. Then $\mathcal{S}$ is semi-globally practically uniformly asymptotically stable with respect to $\mathcal{I}$ for (3).

When dealing with saddle point algorithms in convex optimization, the Lagrange multipliers λ_i are part of the state variables. Their vector field is designed in such a way that the positive orthant is positively invariant for the Lagrange multipliers. The set $\mathcal{I}$ will be the set which is uniformly positively invariant for (3) and which restricts the initial conditions. This is the main difference of Lemma 1 to Theorem 3 in Dürr et al. [2013].

As already used in Lemma 1 we introduce a notion of stability which differs slightly from Lyapunov stability. Systems like (3) are characterized by the fact that due to the persistent excitations of the periodic inputs u_i, the trajectories may not converge to the set $\mathcal{S}$ but to a region which can be made arbitrarily small by choosing ω sufficiently large. Hereby, we speak of practical stability which is captured by the definitions below. For this purpose, the a-neighborhood of the set $\mathcal{S}$ with $a \in (0, \infty)$ is denoted by $\mathcal{U}_a^{\mathcal{S}} := \{x \in \mathcal{I} : \inf_{y \in \mathcal{S}} |x - y| < a\}$

Definition 1. $\mathcal{S} \subseteq \mathbb{R}^n$ is said to be **practically uniformly stable with respect to the set** $\mathcal{I}$ for (3) if for every $\epsilon \in (0, \infty)$ there exists a $\delta \in (0, \infty)$ and $\omega_0 \in (0, \infty)$ such that for all $t_0 \in \mathbb{R}$ and for all $\omega \in (\omega_0, \infty)$

$$x(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow x(t) \in \mathcal{U}_\epsilon^{\mathcal{S}}, t \in [t_0, \infty). \tag{6}$$

Definition 2. The solutions of (3) are said to be **practically uniformly bounded with respect to the set** $\mathcal{I}$ if for every $\delta \in (0, \infty)$ there exists an $\epsilon \in (0, \infty)$ and $\omega_0 \in (0, \infty)$ such that for all $t_0 \in \mathbb{R}$ and for all $\omega \in (\omega_0, \infty)$

$$x(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow x(t) \in \mathcal{U}_\epsilon^{\mathcal{S}}, t \in [t_0, \infty). \tag{7}$$

Definition 3. Let $\delta \in (0, \infty)$. $S \subseteq \mathbb{R}^n$ is said to be **practically uniformly attractive with respect to the set** $\mathcal{I}$ for (3) if for every $\delta, \epsilon \in (0, \infty)$ there exists a $t_f \in [0, \infty)$ and ω_0 such that for all $t_0 \in \mathbb{R}$ and all $\omega \in (\omega_0, \infty)$

$$x(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow x(t) \in \mathcal{U}_\epsilon^{\mathcal{S}}, t \in [t_0 + t_f, \infty). \tag{8}$$

Definition 4. $\mathcal{S} \subseteq \mathbb{R}^n$ is said to be **semi-globally practically uniformly asymptotically stable with respect to** $\mathcal{I}$ for (3) if it is practically uniformly stable with respect to $\mathcal{I}$, practically uniformly bounded with respect to $\mathcal{I}$ and practically uniformly attractive with respect to $\mathcal{I}$.

Note that Definition 1 (practical uniform stability with respect to $\mathcal{I}$) implies that the solutions of (3) are uniformly positively invariant with respect to $\mathcal{I}$. This notion is defined as follows.

Definition 5. The solutions of (3) are said to be **uniformly positively invariant with respect to** $\mathcal{I}$ if for all $t_0 \in \mathbb{R}$

$$x(t_0) \in \mathcal{I} \Rightarrow x(t) \in \mathcal{I}, t \in [t_0, \infty). \tag{9}$$

If (3) is not depending on a parameter ω as it is the case for (4), the definitions above coincide with the usual notion of Lyapunov-stability of a set $\mathcal{S}$ with restricted initial conditions $\mathcal{I}$. In this case we drop the terms "practically" and "semi" in the definitions above.

2.3 Saddle Point Algorithms

In this section, we review some continuous-time optimization algorithms for convex optimization problems of the form (1). We refer to Elster [1978], Bertsekas [1995] for further reading on convex optimization. We distinguish between strictly convex and linear optimization problems. The following results were published in Dürr et al. [2011] and Dürr et al. [2012]. We denote the set of saddle points of a Lagrangian L, i.e. the points $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m_+$ which satisfy $L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*)$ for all $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m_+$ by

$$\begin{aligned} S_L = \{(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m_+ :& \\ (x^*, \lambda^*) \text{ saddle point of } L\}.& \end{aligned} \tag{10}$$

In the following theorem, a continuous-time saddle point algorithm for strictly convex optimization problems is considered and stability of S_L is established. It will turn out later, that this algorithm coincides with the associated Lie bracket system of an extremum seeking system.

Theorem 1. Consider (1), L in (2) and

$$\begin{aligned} \dot{x} &= -\Gamma(\nabla_x f(x) + \sum_{i=1}^{m} \lambda_i \nabla_x g_i(x)) \\ &= -\Gamma \nabla_x L(x, \lambda) && (11a) \\ \dot{\lambda}_i &= \lambda_i g_i(x),\ i = 1, \ldots, m, && (11b) \end{aligned}$$

where Γ is a constant, positive definite matrix. Let L be strictly convex in x. Furthermore, let S_L be non-empty and compact. Then S_L is globally uniformly asymptotically stable with respect to $\mathbb{R}^n \times \mathbb{R}^m_{++}$.

The proof goes along the same lines as in Dürr et al. [2012] using the Lyapunov function

$$\begin{aligned} V =& \frac{1}{2}(x - x^*)^\top \Gamma^{-1}(x - x^*) \\ &+ \sum_{i=1}^{m} \lambda_i - \lambda_i^* - \lambda_i^* \log(\lambda_i) + \lambda_i^* \log(\lambda_i^*), \end{aligned} \tag{12}$$

with some $(x^*, \lambda^*) \in S_L$ and the convention $0 \log(0) = 0$.

In the next theorem, we extend this result to linear programs which are a special case of (1).

Theorem 2. Consider (1) and let $f(x) = c^\top x$, $g_i(x) = a_i^\top x - b_i$ with $c, a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, $i = 1, \ldots, m$. Moreover consider L in (2) and

$$\begin{aligned} \dot{x} &= -\Gamma(c + \sum_{i=1}^{m} \lambda_i a_i + \sum_{i=1}^{m} \lambda_i a_i (a_i^\top x - b_i)) && (13a) \\ \dot{\lambda}_i &= \lambda_i (a_i^\top x - b_i), i = 1, \ldots, m, && (13b) \end{aligned}$$

with a constant, positive definite matrix Γ. Furthermore, let S_L be a singleton. Then S_L is globally uniformly asymptotically stable with respect to $\mathbb{R}^n \times \mathbb{R}^m_{++}$.

The proof goes along the same lines as in Dürr et al. [2011] using the Lyapunov function

$$\tilde{V} = V - \sum_{i=1}^{m} \lambda_i (a_i^\top x^* - b_i) \tag{14}$$

with $(x^*, \lambda^*) \in S_L$ and V in (12).

It will turn out in the following, that the differentiability properties on f and g_i, $i = 1, \ldots, m$ are crucial. These functions will appear later in the vector field of the extremum seeking systems. As mentioned above, we consider extremum seeking systems which can be written in the form of (3). In order to be able to apply Lemma 1 the vector fields must satisfy Assumption A1, i.e. they must be twice continuously differentiable. The dynamics of the λ_i's play an important role at this point. One could also imagine to use a similar approach as in Arrow et al. [1958] and Feijer and Paganini [2010], i.e. $\dot{\lambda}_i = \mathcal{P}(\lambda_i, g_i(x))$, with $\mathcal{P}(\lambda_i, g_i(x)) := 0$ if $\lambda_i = 0$ and $g_i(x) < 0$, and $\mathcal{P}(\lambda_i, g_i(x)) := g_i(x)$ otherwise, $i = 1, \ldots, m$. However, this renders the vector field non-smooth and does therefore not satisfy Assumption A1.

3. MAIN RESULTS

In the following we utilize the framework introduced in Dürr et al. [2013] in order to develop a saddle point seeking systems for convex optimization problems. We propose an extremum seeking scheme, whose Lie bracket approximation system coincides with the saddle points algorithms from the foregoing subsection.

3.1 Main Idea

By considering an optimization problem in one variable, i.e., consider (1) and suppose that $x \in \mathbb{R}$, we illustrate the main idea. We extend the extremum seeking feedback which can be found e.g. in Zhang et al. [2007], Dürr et al. [2013] by using L as a nonlinear map and additionally introducing λ in (2) as a separate state. This is illustrated in Fig. 1.

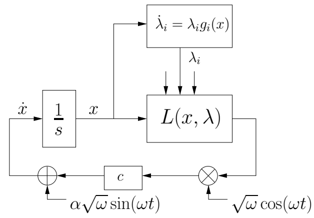


Fig. 1. Saddle Point Seeking for a Convex Optimization Problem in a Single Variable

The extremum seeking system is given by

$$\begin{aligned} \dot{x} &= cL(x, \lambda)\sqrt{\omega}\cos(\omega t) + \alpha\sqrt{\omega}\sin(\omega t) && (15a) \\ \dot{\lambda}_i &= \lambda_i g_i(x), i = 1, \ldots, m. && (15b) \end{aligned}$$

Since the objective function f and the constraints g_i are by assumption physically measurable quantities, they may be part of the vector field of the extremum seeking system. They appear in (15a) in the Lagrangian and in (15b). Note also that due to the multiplication of λ_i in (15b) we immediately see that the solutions of (15) are uniformly positively invariant with respect to the set $\mathbb{R} \times \mathbb{R}^m_{++}$. We write (15) in input-affine form

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda}_1 \\ \vdots \\ \dot{\lambda}_m \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda_1 g_1(x) \\ \vdots \\ \lambda_m g_m(x) \end{bmatrix} + \begin{bmatrix} cL(x, \lambda) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \sqrt{\omega} u_1 + \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix} \sqrt{\omega} u_2 \tag{16}$$

with $u_1 = \cos(\omega t)$ and $u_2 = \sin(\omega t)$. The Lie bracket system associated with (16) is given by

$$\begin{aligned} \dot{\bar{x}} &= -\frac{\alpha c}{2} \nabla_{\bar{x}} L(\bar{x}, \bar{\lambda}) && (17a) \\ \dot{\bar{\lambda}}_i &= \lambda g_i(\bar{x}), i = 1, \ldots, m. && (17b) \end{aligned}$$

We can see two properties of (17). First, the solutions of (17) are uniformly positively invariant with respect to the set $\mathbb{R} \times \mathbb{R}^m_{++}$. This property is inherited from (16). Second, (17) coincides with (11) for $\Gamma = \frac{\alpha c}{2}$.

In order to prove practical uniform asymptotic stability of the set of saddle points S_L of L we must show that

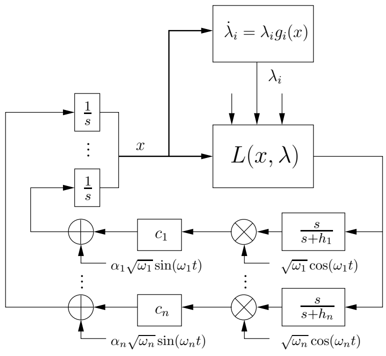


Fig. 2. Saddle Point Seeking for Convex Optimization Problems

S_L is globally uniformly asymptotically stable for (17). Suppose that L is strictly convex and that the set of saddle points S_L is non-empty and compact. We make two observations. First, one can verify that Assumptions A1 – A3 are satisfied for the vector field in (15). Second, since the assumptions of Theorem 1 are satisfied, we can directly conclude that the set of saddle points of L is globally uniformly asymptotically stable for (17) with respect to $\mathbb{R} \times \mathbb{R}^m_{++}$.

Thus, all assumptions of Lemma 1 are satisfied and we conclude that the set of saddle points of L is semi-globally practically uniformly asymptotically stable for (15) with respect to $\mathbb{R} \times \mathbb{R}^m_{++}$.

3.2 Saddle Point Seeking for Convex Optimization Problems

The previous idea is now extended to strictly convex optimization problems in multiple variables. Consider the extremum seeking feedback in Fig. 2. We see that for every component x_i of $x = [x_1, \dots, x_n]^\top$ an extremum seeking feedback is introduced while the sinusoidal perturbations have different frequencies ω_i for each component. Furthermore, in each extremum seeking feedback we see a washout filter with parameter h_i. This filter is common in the extremum seeking literature and it provides a better transient behavior (see e.g. Zhang et al. [2007], Krstić and Ariyur [2003]). It does not influence the stability as we show below in Theorem 3.

The extremum seeking system is given by

$$\dot e_k = - h_k e_k + L(x, \lambda) \tag{18a}$$
$$\dot x_k = c_k(-h_k e_k + L(x,\lambda))\sqrt{\omega_k}\cos(\omega_k t) + \alpha_k\sqrt{\omega_k}\sin(\omega_k t), k = 1, \dots, n \tag{18b}$$
$$\dot\lambda_i = \lambda_i g_i(x), i = 1, \dots, m. \tag{18c}$$

We impose the following assumption on the parameters:

D1 $\omega_k = a_k\omega$ and $a_k \neq a_l, k \neq l$, $a_i \in \mathbb{Q}_{++}$, $\omega \in (0, \infty)$, $h_k, \alpha_k, c_k \in (0,\infty)$, $k, l = 1, \dots, n$.

Next, we define the set

$$E_{S_L} = \left\{ e^* \in \mathbb{R}^n : e^* = L(x^*, \lambda^*)\left[\frac{1}{h_1}, \dots, \frac{1}{h_n}\right]^\top, (x^*, \lambda^*) \in S_L \right\}, \tag{19}$$

which will be used in the following as the set which is practically attractive for the states $e = [e_1, \dots, e_n]^\top$.

Theorem 3. Consider (1), L in (2) and (18). Let L be strictly convex in x. Furthermore, let S_L be non-empty and compact and Assumption D1 be satisfied, then the set $E_{S_L} \times S_L$ is semi-globally practically uniformly asymptotically stable with respect to $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m_{++}$.

Proof. The proof consists of three steps and goes along the procedure introduced in Dürr et al. [2013].

First, we calculate the corresponding Lie bracket system for (18). Similar as in the proof of Theorem 4 in Dürr et al. [2013] we make use of Assumption D1. We obtain the Lie bracket system

$$\dot{\bar e}_k = -h_k \bar e_k + L(\bar x, \bar\lambda), k = 1, \dots, n \tag{20a}$$
$$\dot{\bar x} = -\Gamma\nabla_{\bar x} L(\bar x, \bar\lambda) \tag{20b}$$
$$\dot{\bar\lambda}_i = \bar\lambda_i g_i(\bar x), i = 1, \dots, m \tag{20c}$$

with $\Gamma = \operatorname{diag}(\frac{\alpha_1 c_1}{2}, \dots, \frac{\alpha_n c_n}{2})$ which is positive definite due to Assumption D1.

Second, we show that (20) is globally uniformly asymptotically stable with respect to $\mathcal{I}$. Note that the subsystem $(\bar x, \bar\lambda)$ in (20) is independent of $\bar e = [\bar e_1, \dots, \bar e_n]^\top$. Furthermore, the subsystem satisfies all assumptions of Theorem 1. Thus, we conclude that S_L is globally uniformly asymptotically stable with respect to $\mathcal{I}$ for the subsystem $(\bar x, \bar\lambda)$. Due to Assumption D1 which yields that $h_k \in (0, \infty)$, the subsystems $\dot{\bar e}_k = -h_k\bar e_k + L(\bar x, \bar\lambda), k = 1, \dots, n$ with $u = L(\bar x, \bar\lambda)$ are linear, ordinary differential equations with exponentially stable origin. Thus, $L(\bar x, \bar\lambda)$ is bounded and therefore $\bar e_k$ is bounded with gain $\frac{1}{h_k}$, $k = 1, \dots, n$. We conclude that the set $E_{S_L} \times S_L$ is globally uniformly asymptotically stable with respect to $\mathcal{I}$.

Third, the extremum seeking system in (18) is uniformly positively invariant with respect to $\mathcal{I}$. Thus, all assumptions of Lemma 1 are satisfied which results in the claim of the theorem. □

In the proof of the previous theorem, it is crucial that the vector field of the subsystem consisting of $(\bar x, \bar\lambda)$ in the Lie bracket system (20) coincides with the vector field of (11). In the following, we exploit this observation and extend the result to linear programs. The idea is to construct an extremum seeking system whose respective Lie bracket system coincides with (13).

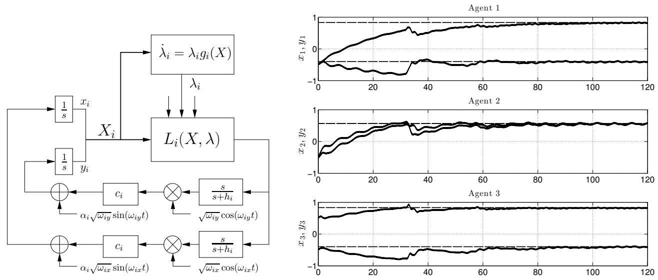
Consider L in (2). One can verify, that the vector field (13a) is the gradient of

$$\tilde L(x, \lambda) = L(x, \lambda) + \frac{1}{2}\sum_{i=1}^{m} \lambda_i (a_i^\top x - b_i)^2, \tag{21}$$

i.e. (13a) can be written as $\dot x = -\nabla_x \tilde L(x, \lambda)$ with $\tilde L$ as defined in (21). Thus, by replacing $L(x, \lambda)$ in the extremum seeking feedback in Fig. 2 with $\tilde L(x, \lambda)$ we obtain

$$\dot e_k = - h_k e_k + \tilde L(x, \lambda) \tag{22a}$$
$$\dot x_k = c_k(-h_k e_k + \tilde L(x,\lambda))\sqrt{\omega_k}\cos(\omega_k t) + \alpha_k\sqrt{\omega_k}\sin(\omega_k t), k = 1, \dots, n \tag{22b}$$
$$\dot\lambda_i = \lambda_i(a_i^\top x - b_i), i = 1, \dots, m. \tag{22c}$$

Theorem 4. Consider (1) and let $f(x) = c^\top x$, $g_i(x) = a_i^\top x - b_i$ with $c, a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, $i = 1, \dots, m$. Moreover consider L in (2) and (22). Furthermore, let S_L be a singleton and let Assumption D1 be satisfied. Then the set $E_{S_L} \times S_L$ is semi-globally practically uniformly asymptotically stable with respect to $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m_{++}$.



(a) (b)

Fig. 3. Distributed Saddle Point Seeking

One can verify that the corresponding Lie bracket system of (22) yields

$$\dot{\bar{e}}_k = -h_k\bar{e}_k + \tilde{L}(\bar{x},\bar{\lambda}), k = 1,\ldots,n \tag{23a}$$

$$\dot{\bar{x}} = -\Gamma\nabla_z\tilde{L}(\bar{x},\bar{\lambda}) \tag{23b}$$

$$\dot{\bar{\lambda}}_i = \bar{\lambda}_i g_i(\bar{x}), i = 1,\ldots,m \tag{23c}$$

with $\Gamma = \mathrm{diag}(\frac{\alpha_1 c_1}{2},\ldots,\frac{\alpha_n c_n}{2})$ which is positive definite due to Assumption D1. Thus, the subsystem consisting of $(\bar{x},\bar{\lambda})$ in the Lie bracket system (23) coincides with the vector field of (13) and with Theorem 2 we conclude that the saddle point S_L is uniformly asymptotically stable for $(\bar{x},\bar{\lambda})$ with respect to $\mathbb{R}^n\times\mathbb{R}^m_{++}$. The rest of the proof goes along the same lines as the proof of Theorem 3.

4. DISTRIBUTED OPTIMIZATION EXAMPLE

In this section, we use the example in the introduction to illustrate an application of the established results above for distributed optimization. Similar as in Dürr et al. [2013] we exploit the fact that the problem can be formulated as a separable optimization problem.

Consider three agents $1,2,3$ with positions $X_1 = [x_1,y_1]^\top$, $X_2 = [x_2,y_2]^\top$, $X_3 = [x_3,y_3]^\top$. We impose distance constraints between agents 1 and 2 as well as between agents 2 and 3 while every agent has its own base station A, B, C. The positions of the base stations are denoted by $X_A = [x_A,y_A]^\top$, $X_B = [x_B,y_B]^\top$, $X_C = [x_C,y_C]^\top$. We formulate this problem as a quadratic program with quadratic constraints

$$\begin{aligned} &\min\ |X_1-X_A|^2+|X_2-X_B|^2+|X_3-X_C|^2\\ &\text{s.t. } |X_1-X_2|^2\le d_1,\\ &\qquad |X_3-X_2|^2\le d_2. \end{aligned} \tag{24}$$

Note that in this problem, the objective function is composed of the distances between the agents and their respective base station and the constraints are the distances among certain agents. Thus, these quantities can be measured through sensors, but not their gradients. The Lagrangian of this problem is given by

$$\begin{aligned} L(X,\lambda) = {}&|X_1-X_A|^2+|X_2-X_B|^2+|X_3-X_C|^2\\ &+\lambda_1(|X_1-X_2|^2-d_1)+\lambda_2(|X_3-X_2|^2-d_2). \end{aligned} \tag{25}$$

Note that the Lagrangian above depends on the position of all agents, i.e. every agent must know the position of every other agent if we considered agents as in Fig. 2. In order to solve the problem in a distributed way, we propose the extremum seeking feedbacks given in Fig. 3a where each agent implements an individual L_i which depends only on its position and the position of neighboring agents.

Since we have $\nabla_{X_i}L_i = \nabla_{X_i}L$ with

$$\begin{aligned} L_1(X,\lambda) &= |X_1-X_A|^2+\lambda_1(|X_1-X_2|^2-d_1)\\ L_2(X,\lambda) &= |X_2-X_B|^2+\lambda_1(|X_1-X_2|^2-d_1)\\ &\quad+\lambda_2(|X_3-X_2|^2-d_2)\\ L_3(X,\lambda) &= |X_3-X_C|^2+\lambda_2(|X_3-X_2|^2-d_2), \end{aligned} \tag{26}$$

it suffices that every agent uses the individual Lagrangian L_i. One can verify that the Lie bracket system corresponding to the overall system is the saddle point system (11) with Lagrangian (25). We see furthermore in (26) that the Lagrange multipliers λ_1 and λ_2 in (25) must be known only to neighboring agents. For the parameters, we choose $X_A = [1,-1]^\top$, $X_B = [1,1]^\top$, $X_C = [-1,1]^\top$, $d_1 = d_2 = 1$, $\alpha_i = c_i = 0.2$, $h_i = 1$, $i = 1,\ldots,3$, $\omega_{1x} = 80, \omega_{1y} = 81, \omega_{2x} = 82,\ldots,\omega_{3y} = 85$. A simulation with these values is shown in Fig. 3b and one can see that the agents converge to the saddle point of L in (25) (the exact solutions are denoted as dashed lines).

5. SUMMARY AND OUTLOOK

We proposed an extremum seeking feedback for saddle points problems arising in convex optimization. Using a Lie bracket approximation, we show that the extremum seeking system can be approximated with a saddle point algorithm. Since the set of saddle points of the Lagrangian associated to the optimization problem is asymptotically stable for the saddle point algorithm, it is practically uniformly asymptotically stable for the extremum seeking system. We also showed that the proposed feedback can be applied to distributed optimization problems with constraints.

It is of future research to generalize the ideas in this paper to more complex dynamics like unicycle models. Another interesting aspect is to consider the case where only the value of the Lagrangian is available. In this case, one can introduce extremum seeking loops for the λ_i's.

Appendix A. PROOF OF LEMMA 1

Note that Assumptions A1 – A3 imply the satisfaction of Assumptions A1 – A4 in Dürr et al. [2013]. Thus we can use Theorem 1 in Dürr et al. [2013]. The rest of the proof follows the same lines as the proof of Theorem 1 in Moreau and Aeyels [2000] except that we now introduce the set $\mathcal{I}$ which is assumed to be uniformly positively invariant and restricts the set of initial conditions.

Practical uniform stability We now show that $\mathcal{S}$ is practically uniformly stable for (3), see Definition 1. Take an arbitrary $\epsilon\in(0,\infty)$ and let $C_1\in(0,\epsilon)$. First observe that, since $\mathcal{S}$ is uniformly stable with respect to $\mathcal{I}$ for (4), there exists a $\delta\in(0,\infty)$ such that for all $t_0\in\mathbb{R}$

$$\bar{x}(t_0)\in\mathcal{U}^{\mathcal{S}}_\delta \Rightarrow \bar{x}(t)\in\mathcal{U}^{\mathcal{S}}_{C_1},\ t\in[t_0,\infty). \tag{A.1}$$

Second observe that, since the set $\mathcal{S}$ is semi-globally uniformly attractive with respect to $\mathcal{I}$ for (4) and we have that for every $C_2\in(0,\delta)$ there exists a time $t_f\in(0,\infty)$ such that for all $t_0\in\mathbb{R}$

$$\bar{x}(t_0)\in\mathcal{U}^{\mathcal{S}}_\delta \Rightarrow \bar{x}(t)\in\mathcal{U}^{\mathcal{S}}_{C_2},\ t\in[t_0+t_f,\infty). \tag{A.2}$$

Let $D = \min\{\epsilon-C_1,\delta-C_2\}$, $\mathcal{B} = \mathcal{K} = \mathcal{U}^{\mathcal{S}}_\delta$ and t_f determined above. Due to Theorem 1 in Dürr et al. [2013], there exists an $\omega_0\in(0,\infty)$ such that for all $\omega\in(\omega_0,\infty)$ and all $x(t_0)\in\mathcal{K}=\mathcal{U}^{\mathcal{S}}_\delta$, $|x(t)-\bar{x}(t)|<D$, $t\in[t_0,t_0+t_f]$. This together with (A.1), (A.2) and the uniform positive invariance of $x(t)$ with respect to $\mathcal{I}$ yield for all $\omega\in(\omega_0,\infty)$

$$x(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow x(t) \in \mathcal{U}_\epsilon^{\mathcal{S}}, t \in [t_0, t_0 + t_f] \text{ and } x(t_0 + t_f) \in \mathcal{U}_\delta^{\mathcal{S}}. \quad \text{(A.3)}$$

Since $x(t_0 + t_f) \in \mathcal{U}_\delta^{\mathcal{S}}$ and a repeated application of the procedure with another solution $\bar{x}(t)$ of (4) through $x(t_0 + t_f)$ with the same choice of D, $\mathcal{K}$ and t_f as above yields for all $t_0 \in \mathbb{R}$, for all $\omega \in (\omega_0, \infty)$

$$x(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow x(t) \in \mathcal{U}_\epsilon^{\mathcal{S}}, t \in [t_0, \infty). \quad \text{(A.4)}$$

Practical uniform boundedness Take an arbitrary $\delta \in (0, \infty)$ and let $C_1 \in (0, \delta)$. Since $\mathcal{S}$ is uniformly bounded with respect to $\mathcal{I}$ and uniformly attractive with respect to $\mathcal{I}$ for (4) there exist $C_2 \in (0, \infty)$ and $t_f \in (0, \infty)$ such that for all $t_0 \in \mathbb{R}$

$$\bar{x}(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow \bar{x}(t) \in \mathcal{U}_{C_2}^{\mathcal{S}},\ t \in [t_0, \infty) \text{ and } \bar{x}(t) \in \mathcal{U}_{C_1}^{\mathcal{S}},\ t \in [t_0 + t_f, \infty). \quad \text{(A.5)}$$

Let $\epsilon \in (C_2, \infty)$, $D = \min\{\delta - C_1, \epsilon - C_2\}$, $\mathcal{B} = \mathcal{K} = \mathcal{U}_\delta^{\mathcal{S}}$ and t_f determined above. Due to Theorem 1 in Dürr et al. [2013], there exists an $\omega_0 \in (0, \infty)$ such that for all $\omega \in (\omega_0, \infty)$ and all $x(t_0) \in \mathcal{K} = \mathcal{U}_\delta^{\mathcal{S}}$, $|x(t) - \bar{x}(t)| < D$, $t \in [t_0, t_0 + t_f]$. This together with (A.5) and the uniform positive invariance of $x(t)$ with respect to $\mathcal{I}$ yield for all $\omega \in (\omega_{0,1}, \infty)$

$$x(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow x(t) \in \mathcal{U}_\epsilon^{\mathcal{S}}, t \in [t_0, t_0 + t_f] \text{ and } x(t_0 + t_f) \in \mathcal{U}_\delta^{\mathcal{S}}. \quad \text{(A.6)}$$

Since $x(t_0 + t_f) \in \mathcal{U}_\delta^{\mathcal{S}}$ and a repeated application of the procedure with another solution $\bar{x}(t)$ of (4) through $x(t + t_f)$ and the same choice of D, $\mathcal{K}$ and t_f as above yields for all $t_0 \in \mathbb{R}$, for all $\omega \in (\omega_0, \infty)$

$$x(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow x(t) \in \mathcal{U}_\epsilon^{\mathcal{S}}, t \in [t_0, \infty). \quad \text{(A.7)}$$

Practical uniform attractivity Choose some $\delta, \epsilon \in (0, \infty)$. By practical uniform stability proven above, there exist $C_1 \in (0, \infty)$ and $\omega_{0,1} \in (0, \infty)$ such that for all $t_0 \in \mathbb{R}$ and for all $\omega \in (\omega_{0,1}, \infty)$

$$x(t_0) \in \mathcal{U}_{C_1}^{\mathcal{S}} \Rightarrow x(t) \in \mathcal{U}_\epsilon^{\mathcal{S}}, t \in [t_0, \infty). \quad \text{(A.8)}$$

Let $\epsilon_1 \in (0, C_1)$. Since the set $\mathcal{S}$ is uniformly attractive for (4), there exists a $t_f \in (0, \infty)$ such that for all $t_0 \in \mathbb{R}$

$$\bar{x}(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow \bar{x}(t) \in \mathcal{U}_{\epsilon_1}^{\mathcal{S}}, t \in [t_0 + t_f, \infty). \quad \text{(A.9)}$$

Note that by uniform boundedness there exists an $A \in (0, \infty)$ such that for every $t_0 \in \mathbb{R}$ we have that $\bar{x}(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow \bar{x}(t) \in \mathcal{U}_A^{\mathcal{S}}, t \in [t_0, \infty)$. Due to Theorem 1 in Dürr et al. [2013] with $\mathcal{B} = \mathcal{K} = \mathcal{U}_\delta^{\mathcal{S}}$, $D = C_1 - \epsilon_1$ and t_f defined above there exists an $\omega_{0,2} \in (0, \infty)$ such that for all $t_0 \in \mathbb{R}$ and for all $\omega \in (\omega_{0,2}, \infty)$ and all $x(t_0) \in \mathcal{K} = \mathcal{U}_\delta^{\mathcal{S}}$ we have that $|x(t) - \bar{x}(t)| < D,\ t \in [t_0, t_0 + t_f]$. This estimate together with (A.9) and the uniform positive invariance of $x(t)$ with respect to $\mathcal{I}$ yield for all $t_0 \in \mathbb{R}$ and for all $\omega \in (\omega_{0,2}, \infty)$

$$x(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow x(t_0 + t_f) \in \mathcal{U}_{C_1}^{\mathcal{S}}. \quad \text{(A.10)}$$

With (A.8), this leads for all $t_0 \in \mathbb{R}$ and for all $\omega \in (\omega_0, \infty)$ with $\omega_0 = \max\{\omega_{0,1}, \omega_{0,2}\}$ to

$$x(t_0) \in \mathcal{U}_\delta^{\mathcal{S}} \Rightarrow x(t) \in \mathcal{U}_\epsilon^{\mathcal{S}}, t \in [t_0 + t_f, \infty). \quad \text{(A.11)}$$

This is the last property we had to prove.

REFERENCES

K. J. Arrow, L. Hurwicz, and H. Uzawa. *Studies in Linear and Non-linear Programming*. Stanford University Press, 1958.

D. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1995.

F. Brunner, H. B. Dürr, and C. Ebenbauer. Feedback Design for Multi-Agent Systems: A Saddle Point Approach. In *Proceedings of the 51st Conference on Decision and Control*, pages 3783 – 3789, Maui, Hawaii, 2012.

F. Coito, J. M. Lemos, and S. S. Alves. Stochastic extremum seeking in the presence of constraints. In *Proceedings of the 16th IFAC World Congress*, 2005.

D. DeHaan and M. Guay. Extremum-seeking control of state-constrained nonlinear systems. *Automatica*, 41(9): 1567 – 1574, 2005.

H. B. Dürr and C. Ebenbauer. A Smooth Vector Field for Saddle Point Problems. In *Proceedings of the 50th IEEE Conference on Decision and Control, Orlando, USA*, pages 4654 – 4660, 2011.

H. B. Dürr, S. S. Zeng, and C. Ebenbauer. Ein nichtlineares System zum Lösen von Linearen Programmen und Sattelpunktproblemen. *17. Steirisches Seminar über Regelungstechnik und Prozessautomatisierung, Graz, Austria*, 2011.

H. B. Dürr, E. Saka, and C. Ebenbauer. A smooth vector field for quadratic programming. In *Proceedings of the 51st IEEE Conference on Decision and Control*, pages 2515 – 2520, Maui, Hawaii, 2012.

H. B. Dürr, M. S. Stanković, C. Ebenbauer, and K. H. Johansson. Lie Bracket Approximation of Extremum Seeking Systems. *Automatica*, 49(6):1538 – 1552, 2013.

K. H. Elster. *Nichtlineare Optimierung*. Harri Deutsch, 1978.

D. Feijer and F. Paganini. Stability of primal-dual gradient dynamics and applications to network optimization. *Automatica*, 46(12):1974 – 1981, 2010.

P. Frihauf, M. Krstić, and T. Bašar. Nash equilibrium seeking in noncooperative games. *IEEE Transactions on Automatic Control*, 57(5):1192 – 1207, 2012.

M. Guay, D. Dochain, and M. Perrier. Adaptive extremum seeking control of continuous stirred tank bioreactors with unknown growth kinetics. *Automatica*, 40(5):881 – 888, 2004.

R. King, R. Becker, G. Feuerbach, L. Henning, R. Petz, W. Nitsche, O. Lemke, and W. Neise. Adaptive flow control using slope seeking. In *14th Mediterranean Conference on Control and Automation*, pages 1 – 6, 2006.

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

L. Moreau and D. Aeyels. Practical stability and stabilization. *IEEE Transactions on Automatic Control*, 45 (8):1554 – 1558, 2000.

A. R. Nazemi. A dynamic system model for solving convex nonlinear optimization problems. *Communications in Nonlinear Science and Numerical Simulation*, 17(4): 1696 – 1705, 2012.

A. Nedić and A. Ozdaglar. Subgradient Methods for Saddle-Point Problems. *Journal of Optimization Theory and Applications*, 142(1):205 – 228, 2009.

J. Poveda and N. Quijano. A Shahshahani gradient based extremum seeking scheme. In *Proceedings of the IEEE International Conference on Decision and Control*, pages 5104 – 5109, Maui, Hawaii, 2012.

Y. Tan, W. H. Moase, C. Manzie, D. Nesić, and I.M.Y. Mareels. Extremum seeking from 1922 to 2010. In *Proceedings of the 29th Chinese Control Conference*, pages 14 – 26, 2010.

C. Zhang, A. Siranosian, and M. Krstić. Extremum seeking for moderately unstable systems and for autonomous vehicle target tracking without position measurements. *Automatica*, 43(10):1832 – 1839, 2007.