Robust Time-Varying Continuous-Time Optimization with Pre-Defined Finite-Time Stability

Romero, Orlando; Benosman, Mouhacine

TR2020-120 August 28, 2020

Abstract

In this paper we propose a new family of continuous-time optimization algorithms based on discontinuous second order gradient optimization flows, with finite-time convergence guarantees to local optima, for locally strongly convex time-varying cost functions. To analyze our flows, we first extend a well-know Lyapunov inequality condition for finite-time stability, to the case of time-varying differential inclusions. We then prove the convergence of these secondorder flows in finite-time. We show the performance of these flows on a time-varying quadratic cost and on the nonlinear time-varying Rosenbrock function.

World Congress of the International Federation of Automatic Control (IFAC)

Robust Time-Varying Continuous-Time Optimization with Pre-Defined Finite-Time Stability

O. Romero *

* *Department of Industrial and Systems Engineering Rensselaer Polytechnic Institute Troy, NY 12180, USA (rodrio2@rpi.edu)*

M. Benosman **

** *Mitsubishi Electric Research Laboratories Cambridge, MA 02139, USA (benosman@merl.com)*

Abstract: In this paper we propose a new family of continuous-time optimization algorithms based on discontinuous second order gradient optimization flows, with finite-time convergence guarantees to local optima, for locally strongly convex *time-varying cost functions*. To analyze our flows, we first extend a well-know Lyapunov inequality condition for finite-time stability, to the case of time-varying differential inclusions. We then prove the convergence of these second-order flows in finite-time. We show the performance of these flows on a time-varying quadratic cost and on the nonlinear time-varying Rosenbrock function.

1. INTRODUCTION

An important class of continuous optimization algorithms are the so-called extremum seeking (ES) controllers, which deal with static cost functions, as well as dynamic cost functions, modeled as the output of a dynamical system. Most importantly, ES algorithms are often based only on the cost function measurements, i.e., zero-order optimization methods, whereas the higher order derivatives of the cost function, e.g., gradient and Hessian, are estimated from the cost function measurements using feedback filters, e.g., Poveda & Teel [2017], Guay& Zhang [2003], Krstić [2000], Ariyur & Krstić [2003], Scheinker & Krstić [2016], Zhang & Ordónez [2012], Grushkovskaya et al. [2018]. Since we are not considering zero-order methods in this work, we will not discuss specifically ES results, and will focus on the more general class of continuous-optimization algorithms, including higher order methods.

For instance, in Su et al. [2016], the authors derive a second-order ODE as the limit of Nesterov's accelerated gradient method, when the gradient step sizes go to zero. This ODE is then used to attempt to analyze Nesterov's scheme, particularly in an larger effort to better understand acceleration without substantially increasing computational burden. Thanks to the ODE continuous-time approximation of the algorithm, the authors also obtain a family of schemes with similar convergence rates as Nestrov's algorithm.

In Franka et al. [2018], The differential equations that model the continuous-time limit of the sequence of iterates generated by the alternating direction method of multipliers (ADMM), are derived. Then, the authors employ Lyapunov theory to analyze the stability of critical points of the dynamical systems and to obtain associated convergence rates.

In Franca et al. [2019a], non-smooth and linearly constrained optimization problems are analyzed by deriving equivalent (at the limit) non-smooth dynamical systems related to variants of the relaxed and accelerated ADMM. In particular, two new ADMM-like algorithms are proposed, one based on Nesterov's acceleration and the other inspired by Polyak's heavy ball method, and derive differential inclusions modeling these algorithms in the continuous-time limit. Using a non-smooth Lyapunov analysis, results on rate-of-convergence are obtained for these dynamical systems in the convex and strongly convex setting.

In Cortes [2006], two normalized first-order gradient flows are proposed. Their convergence is rigorously analyzed using tools from non-smooth dynamics theory, and conditions guaranteeing finite-time convergence are derived. Finally, the proposed non-smooth flows are applied to problems in multi-agent systems and it is shown they achieve consensus in a finite-time. The finite convergence time's upper bound is given as function of the gradient value at the initial point as well as the minimum eigenvalue of Hessian at the initial point.

More recently, in Poveda & Li [2019], the authors establish uniform asymptotic stability and robustness properties for the continuous-time limit of the Nestrov's accelerated gradient method, and in Yuan et al. [2019] the authors propose a powerball method to accelerate the convergence of some first order and second order optimization algorithms for static cost functions.

In Romero & Benosman [2019], the authors introduce two discontinuous dynamical systems in continuous time with guaranteed prescribed finite-time local convergence to strict local minima of a given static cost function. In this work, we extend these results to the case of time-varying costs, and propose a new family of *discontinuous second-order flows*, which guarantee local convergence to an optimum, in a *desired pre-defined finite-time, for time-varying cost functions*. We use some ideas from Lyapunov-based finite-time state control to an invariant set, proposed by one of the current authors in an early paper Benosman & Lum [2009], in the context of aerospace applications, to design a new family of discontinuous flows, which ensure a desired finite-time convergence to the invariant set containing a unique local optima. Furthermore, due to the discontinuous nature of the proposed flows, we propose to extend one of the existing Lyapunov-based inequality condition for finite-time convergence of continuous-time dynamical systems, to the case of time-varying differential inclusions.

This paper is organized as follows: Section 2 is dedicated to some preliminaries about continuous-time optimization, and finite-time stability in the context of time-varying differential inclusions. Our main results are presented in Section 3, where we first establish an extension to time-varying differential inclusions of a well-know Lyapunov-based inequality condition for finite-time stability. We then propose and analyze our second-order discontinuous flows, including a flow for time-varying cost functions. In Section 4, we show the efficiency of this continuous-time optimization flow on a well established optimization benchmarks. The paper ends with a summarizing conclusion and a discussion of our ongoing investigations, in Section 5.

2. PRELIMINARIES

2.1 Filippov Differential Inclusion for Time-Variant Systems

Similarly to the time-invariant case, a solution to an initial value problem

$$\dot{x}(t) = F(t, x(t)) \tag{1a}$$

$$x(0) = x_0 \tag{1b}$$

with $F : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$ is typically guaranteed to exist and be unique by ensuring that $F(\cdot, x)$ is continuous near $x = x^\star$ and $F(t, \cdot)$ is Lipschitz continuous near $t = 0$. When $F(t, \cdot)$ is not Lipschitz continuous (*e.g.* due to singularities or discontinuities), we understand solutions to (1a) in the sense of Filippov. More precisely, $x : [0, \tau) \to \mathbb{R}^n$ with $0 < \tau \leq \infty$ is a *Filippov solution* to (1) if it is absolutely continuous, $x(0) = x_0$, and

$$\dot{x}(t) \in \mathcal{K}[F](t, x(t)) \tag{2}$$

holds almost everywhere (a.e.) within every compact subinterval of $[0, \tau)$, where $\mathcal{K}[F]$ denotes the Filippov set-valued map Paden & Sastry [1987], Cortes [2008] given by

$$\mathcal{K}[F](t, x) \triangleq \bigcap_{\delta>0} \bigcap_{\mu(S)=0} \overline{\text{co}}\, F(t, B_\delta(x) \setminus S), \tag{3}$$

where μ denotes the Lebesgue measure and $\overline{\text{co}}$ the convex closure. Furthermore, $x(\cdot) : [0, \tau) \to \mathbb{R}^n$ is a *maximal* Filippov solution if it cannot be extended, *i.e.* if no Filippov solution exists over an interval $[0, \tau')$ with $\tau' > \tau$.

Assumption 1. F is Lebesgue measurable and locally essentially bounded, *i.e.* given any (t, x), F is bounded a.e. on every bounded neighborhood of (t, x).

Under Assumption 1, at least one Filippov solution to (1) must exist Paden & Sastry [1987], Cortes [2008]. Furthermore, the Filippov set-valued map (3) can be computed as

$$\mathcal{K}[F](t, x) = \overline{\text{co}} \left\{ \lim_{k\to\infty} F(t, x_k) : \mathcal{N}_F \cup S \not\ni x_k \to x \right\} \tag{4}$$

for some set $\mathcal{N}_F \subset \mathbb{R}^n$ of measure zero and any other set $S \subset \mathbb{R}^n$ of measure zero. In particular, if $F(t, \cdot)$ is continuous at a fixed point x, then $\mathcal{K}[F](t, x) = \{F(t, x)\}$. For instance, for the gradient flow, we have $\mathcal{K}[-\nabla f](t, x) = \{-\nabla f(x)\}$ for every $x \in \mathbb{R}^n$, provided that f is continuously differentiable. Furthermore, if f is only Lipschitz continuous, then $\mathcal{K}[-\nabla f](t, x) = -\partial f(x)$, where ∂f denotes Clarke's generalized gradient Clarke [2001].

2.2 Finite-Time Stability for Time-Variant Differential Inclusions

Consider a general time-varying differential inclusion Bacciotti & Ceragioli [1999]

$$\dot{x}(t) \in K(t, x(t)) \tag{5a}$$

$$x(0) = x_0 \tag{5b}$$

where $K : \mathbb{R}_+ \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is an arbitrary set-valued map.

Assumption 2. $K : \mathbb{R}_+ \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is an upper semi-continuous set-valued map, with nonempty, compact, and convex values.

For instance, in Filippov & Arscott [1988] the authors proved that, under Assumption 1, $K = \mathcal{K}[F]$ satisfies Assumption 2.

We say that $x : [0, \tau) \to \mathbb{R}^n$ with $0 < \tau \leq \infty$ is a *Carathéodory solution* to (5) if $x(\cdot)$ is absolutely continuous on any closed subinterval of $[0, \tau)$, (5a) is satisfied a.e. within every compact subinterval of $[0, \tau)$, and $x(0) = x_0$.

Proposition 1. Under Assumption 2, at least one Carathéodory solution to (5) must exist. In particular, under Assumption 1, at least one Filippov solution to (1) must exist.

We say that $x : [0, \tau) \to \mathbb{R}^n$ is a *maximal* Carathéodory solution of (5) if it cannot be extended, *i.e.* if no solution exists over an interval $[0, \tau')$ with $\tau' > \tau$. In particular, (maximal) Filippov solutions to (1) are nothing but (maximal) Carathéodory solutions to the Filippov differential inclusion (2) with initial condition $x(0) = x_0$.

Furthermore, we say that $x^\star \in \mathbb{R}^n$ is an *equilibrium* of (5) if $x(t) \equiv x^\star$ over $(0, \infty)$ is a Carathéodory solution to (5). In other words, if $0 \in K(t, x^\star)$ holds a.e. in $t \geq 0$. We say that (5) is *(strongly) Lyapunov stable* at $x^\star \in \mathbb{R}^n$ if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that, for every Carathéodory solution $x(\cdot)$ of (5), we have $\|x_0 - x^\star\| < \delta \implies \|x(t) - x^\star\| < \varepsilon$ for every $t \geq 0$ in the interval where $x(\cdot)$ is defined. Furthermore, we say that (5) is *(locally and strongly) asymptotically stable* at $x^\star \in \mathbb{R}^n$ if it is Lyapunov stable at $x^\star$ and there exists some $\delta > 0$ such that every maximal Carathéodory solution $x(\cdot)$ to (5) is defined over $[0, \infty)$ and, if $\|x_0 - x^\star\| < \delta$ then $x(t) \to x^\star$ as $t \to \infty$. Finally, we say that (5) is *(locally and strongly) finite-time stable* at $x^\star \in \mathbb{R}^n$ if it is asymptotically stable at $x^\star$ and there exist some $\delta > 0$ and positive definite function (w.r.t. $x^\star$) $T : B_\delta(x^\star) \to \mathbb{R}_+$ (called the *settling time*) such that, for every Carathéodory solution $x(\cdot)$ of (5) with $x_0 \in B_\delta(x^\star) \setminus \{x^\star\}$, we have $x(t) \in B_\delta(x^\star) \setminus \{x^\star\}$ for every $t \in [0, T(x_0))$ and $x(t) \to x^\star$ as $t \to T(x_0)$.

3. MAIN RESULTS

To establish finite-time stability, first we will propose an extension to the case of time-variant differential inclusions of a well-know Lyapunov-based result for the case of systems of the form (1), with Lipschitz continuous flow $F(.)$, e.g., see (Lemma 1 in Benosman & Lum [2009]). Next, we will use these results to analyze the stability of our discontinuous gradient-like flows for continuous-time optimization.

Theorem 1. Let $x^\star \in \mathbb{R}^n$ be an equilibrium point of (5) and let $V : \mathbb{R}_+ \times \mathcal{D} \to \mathbb{R}$ be a continuously differentiable and positive definite function w.r.t. $x^\star$, where $\mathcal{D} \subset \mathbb{R}^n$ is an open and positively invariant neighborhood of $x^\star$. Suppose that $K(t, x) = \mathcal{K}[F(t, \cdot)](x)$ is nonempty for every $x \in \mathcal{D}$. Let

$$\dot{V}(t,x) \overset{\text{def}}{=} \left\{ \frac{\partial V}{\partial t}(t,x) + \nabla V(t,x) \cdot v : v \in K(t,x) \right\} \quad (6)$$

for $t \geq 0$ and $x \in \mathcal{D}$, where $\nabla V(t,x)$ denotes the gradient of $V(t,x)$ w.r.t x. If there exist constants $c > 0$ and $\alpha \in (0,1)$ such that

$$\sup \dot{V}(t,x) \leq -c[V(t,x)]^{\alpha} \quad (7)$$

a.e. in $t \geq 0$ and $x \in \mathcal{D}$, then $x(t) \to x^{\star}$ in finite-time for every solution $x(\cdot)$ of (5) with $x_0 \in \mathcal{D}$, and the settling time $t^{\star}$ is upper bounded by

$$t^{\star} \leq \frac{V(0,x_0)^{1-\alpha}}{c(1-\alpha)}. \quad (8)$$

Furthermore, if $\dot{V}(t,x)$ contains a single point a.e. in $x \in \mathcal{D}$ and (7) is exact, then so is (8).

Proof 1. Lyapunov stability follows from [Filippov & Arscott, 1988, 3§15 – Theorem 1] for time-varying differential inclusions, which also tells us. that the origin is an equilibrium. Now, given an arbitrary Carathéodory solution $x(\cdot)$ of (5), note that $\mathcal{E}(t) \triangleq V(t,x(t))$ is absolutely continuous (Appendix, Lemma 2) due to V being continuously differentiable. Therefore, since $\frac{d}{dt}V(t,x(t)) = \dot{\mathcal{E}}(t) \in \dot{V}(t,x(t))$ [Bacciotti & Ceragioli, 1999, Lemma 1], we note from (7) that

$$\frac{d}{dt}V(t,x(t)) \leq -cV(t,x(t))^{\alpha}, \quad (9)$$

a.e. in $t \geq 0$. The rest of the proof follows by integrating and setting $x(T(x_0)) = 0$.

Next, we propose to use the result of Theorem 11, to design a discontinuous flow with finite-time convergence guarantees.

Let us first introduce the following assumptions.

Assumption 3. Let $f : \mathbb{R}_{+} \times \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable in both variables, let $x^{opt} : \mathbb{R}_{+} \to \mathcal{D}$ s.t., for each t, $x^{opt}(t)$ be a strict local minimizer (respectively, maximizer) and isolated stationary point of $f(t,\cdot)$, where $\mathcal{D} \subset \mathbb{R}^n$ is an open set s.t. $x^{opt}(t) \in \mathcal{D}$, $\forall t \geq 0$. Then, we assume that $\nabla^2 f(t,x) < 0$ (respectively > 0), $\forall x \in \mathcal{D}$, $\forall t \geq 0$.

Assumption 4. Let $f : \mathbb{R}_{+} \times \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable, in both variables, let $x^{opt} : \mathbb{R}_{+} \to \mathcal{D}$ s.t., for each t, $x^{opt}(t)$ be a strict local optima and isolated stationary point of $f(t,\cdot)$, where $\mathcal{D} \subset \mathbb{R}^n$ is an open set s.t. $x^{opt}(t) \in \mathcal{D}$, $\forall t \geq 0$. Then, we assume the existence of a continuous function $l : \mathbb{R}_{+} \times \mathbb{R}^n \to \mathbb{R}$, such that

$$\| \frac{\partial}{\partial t}[\nabla f(t,x)] \| \leq l(t,x), \forall t \geq 0, \forall x \in \mathcal{D}. \quad (10)$$

We can now present the main result of this paper.

Proposition 2. Let $f : \mathbb{R}_{+} \times \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable, in both variables, let $x^{opt} : \mathbb{R}_{+} \to \mathcal{D}$ s.t., for each t, $x^{opt}(t)$ be a strict local minimizer (respectively, maximizer) and isolated stationary point of $f(t,\cdot)$, where $\mathcal{D} \subset \mathbb{R}^n$ is an open set s.t. $x^{opt}(t) \in \mathcal{D}$, $\forall t \geq 0$. Consider the flow given by

$$\dot{x} = -\frac{1}{2} \frac{[\nabla^2 f(t,x)]^r \nabla f(t,x)}{\nabla f(t,x)^T [\nabla^2 f(t,x)]^{r+1} \nabla f(t,x)} \big(2l(t,x) \| \nabla f(t,x) \| + c \| \nabla f(t,x) \|^{2\alpha} \big), \quad (11)$$

with $c > 0$, $\alpha \in [0.5, 1)$, $r \in \mathbb{R}$, and where $l : \mathbb{R}_{+} \times \mathbb{R}^n \to \mathbb{R}$ satisfies Assumption 4. Then, under Assumption 3, any Filippov solution $x(\cdot)$ of (11), with $x(0) = x_0$ sufficiently close to $x^{opt}(t)$ for a given $t \geq 0$, will converge in finite-time to $x^{opt}(.)$ with a settling time $t^{\star} \leq \frac{\|\nabla f(0,x_0)\|^{2(1-\alpha)}}{c(1-\alpha)}$.

Proof 2. Let us define the tracking error as $e = x - x^{opt}(t)$, we then consider the Lyapunov function $V(t,e) = \|\nabla f(t, e + x^{opt}(t))\|^2$, and write its derivative as follows, for $e \in \{x - x^{opt} : x \in \mathcal{D}\} \setminus \{0\}$:

$$\begin{aligned}
\sup \dot{V}(t,e) &= \frac{\partial}{\partial t}[\nabla f(t,e + x^{opt}(t))^T \nabla f(t,e + x^{opt}(t))] \\
&\frac{\partial}{\partial e}[\nabla f(t,e + x^{opt}(t))^T \nabla f(t,e + x^{opt}(t))]\dot{e}, \\
&= \frac{\partial}{\partial t}[\nabla f(t,e + x^{opt}(t))^T \nabla f(t,e + x^{opt}(t))] \\
&\quad +2\nabla f(t,x)^T[\nabla^2 f(t,x)](\dot{x} - \dot{x}^*(t)), \\
&= \frac{\partial}{\partial t}[\nabla f(t,x)^T \nabla f(t,x)] + \frac{\partial}{\partial x}[\nabla f(t,x)^T \nabla f(t,x)]\dot{x}^*(t) \\
&\quad +2\nabla f(t,x)^T[\nabla^2 f(t,x)](\dot{x} - \dot{x}^*(t)) \\
&= \frac{\partial}{\partial t}[\nabla f(t,x)^T \nabla f(t,x)] + 2\nabla f(t,x)^T[\nabla^2 f(t,x)]\dot{x},
\end{aligned} \quad (12)$$

next, by using (11), we can write

$$\begin{aligned}
\sup \dot{V}(t,e) &= \frac{\partial}{\partial t}[\nabla f(t,x)^T \nabla f(t,x)] - 2l(t,x)\|\nabla f(t,x)\| \\
&\quad -c\|\nabla f(t,x)\|^{2\alpha} \\
&\leq -c\|\nabla f(t,e + x^{opt}(t)\|^{2\alpha} = -cV(t,e)^{\alpha},
\end{aligned} \quad (13)$$

which, by Theorem 1, leads to the desired finite-time convergence result.

Remark 1. (Robustness). It is clear from equation (13) that if instead of using an upper-bound $l(t,x)$ of the norm of $\frac{\partial}{\partial t}[\nabla f(t,x)]$ in the flow (11), one uses the exact term $-\frac{\partial}{\partial t}[\nabla f(t,x)^T \nabla f(t,x)]$, we can obtain an exact value of the finite-time convergence, i.e., $t^* = \frac{\|\nabla f(0,x_0)\|^{2(1-\alpha)}}{c(1-\alpha)}$. However, this will not be very practical, since it is difficult to be able to obtain the term $\frac{\partial}{\partial t}[\nabla f(t,x)^T \nabla f(t,x)]$ in closed-form is any meaningful application, and its numerical approximation will induce numerical errors, implying a lack of robustness of this solution, since it is based on an exact cancellation of this time-varying term.

4. NUMERICAL EXAMPLES

First, we consider the quadratic time-varying cost function

$$f(t,x) = (x - (8 + 4sin(t)))^2. \quad (14)$$

We apply the flow given by (11), with different upper-bound functions l. We chose the constants to be $r = -1$, $\alpha = 0.5$, $c = \|\nabla f(0, x(0))\|$, and the initial condition $x(0) = 4$. We report below the numerical results for each upper-bound function l. In the first case, we assume exact knowledge of the cost function and compute the upper-bound in (10) in closed-form as $l(t,x) = 32(12 + \|x\|)$. The corresponding results are reported in Figure 1. It is clear that the flow manages to minimize the time-varying cost function in finite-time. To appreciate this result, we also report in Figure 2 the results when we force l to zero. One can see that the flow without the upper-bound term l does not converge, which underlines the necessity of this term in ensuring convergence, as see in the analysis of the results of Proposition 2.

In the previous result we used an exact upper-bound l derived from closed-from manipulation of the cost function.

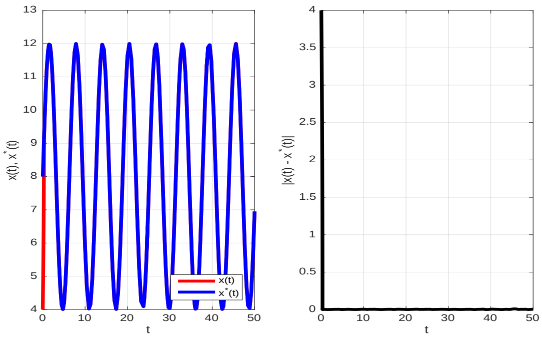

Fig. 1. Time-varying cost function: Flow (11) with $l(t, x) = 32(12 + \|x\|)$.

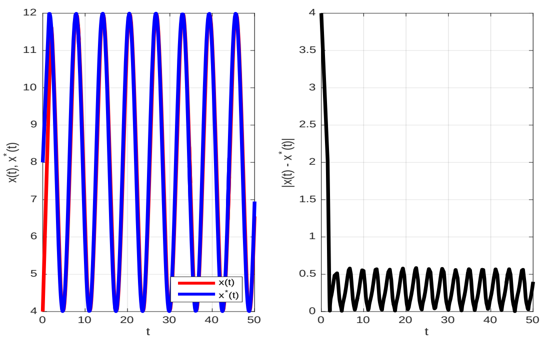

Fig. 2. Time-varying cost function: Flow (11) with $l(t, x) = 0$.

However, this is not always possible in real applications, and thus we show next that any loose upper-bound suffice to ensure the finite-time stability result. Indeed, we first report in Figure 5 the results obtained with the positive definite function $l(t, x) = 50+t^2$. Then we report in Figure 6 the results corresponding to the case $l(t, x) = 100$, which is the simplest upper-bound one can choose. In both cases the flow achieves the expected finite-time convergence. This shows that the proposed flow (11) is not very sensitive to the choice of the function l, as long as it is a valid upper-bound, as defined in (10).

Finally, we test the performance of our flow on the more challenging time-varying Rosenbrock cost function

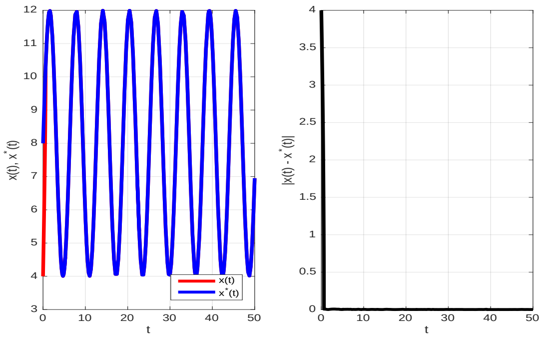

Fig. 3. Time-varying cost function: Flow (11) with $l(t, x) = 50 + t^2$.

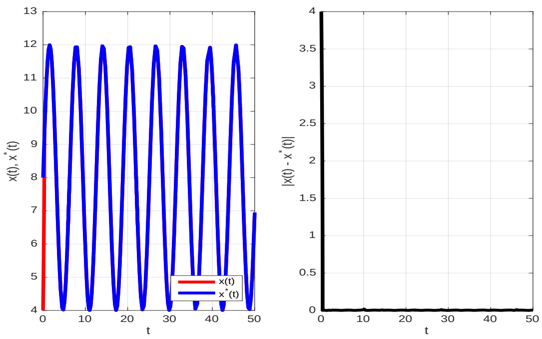

Fig. 4. Time-varying cost function: Flow (11) with $l(t, x) = 100$.

$$\begin{aligned} &f(t, x_1, x_2) = (a(t) - x_1)^2 + b(t)(x_2 - x_1^2)^2, \\ &a(t) = 2 + sin(t), \\ &b(t) = 50(1 + sin(t)), \end{aligned} \tag{15}$$

We apply the flow given by (11), with different upper-bound functions l. We chose the constants to be $r = -1,\ \alpha = 2/3,\ c = 100$, and the initial condition $x(0) = (4,\ 15)'$. we first report in Figure 5 the results obtained with the positive definite function $l(t, x) = 500 + t^2$. Then we report in Figure 6 the results corresponding to the case $l(t, x) = 500$, which is the simplest upper-bound one can choose. In both cases the flow achieves the expected finite-time convergence. This shows that the proposed flow (11) is robust to the choice of the function l, as long as it is a valid upper-bound, as defined in (10).

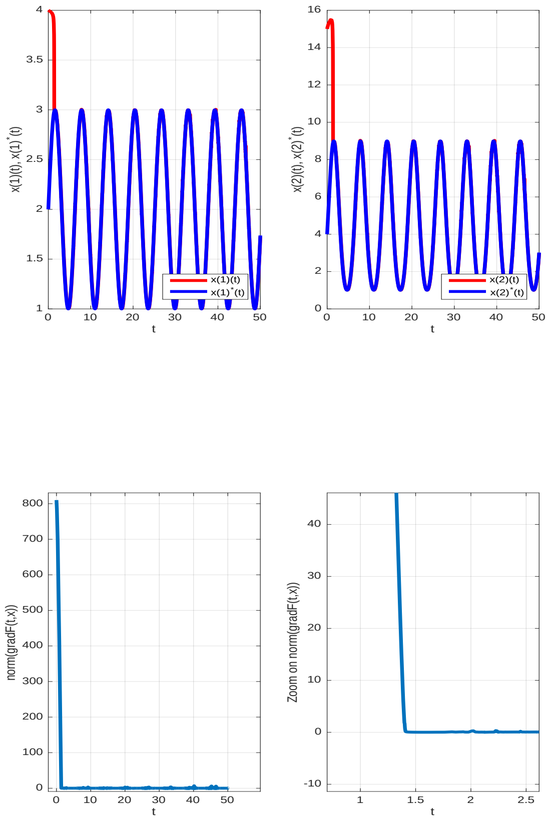

Fig. 5. Time-varying cost function: Flow (11) with $l(t,x) = 500 + t^2$.

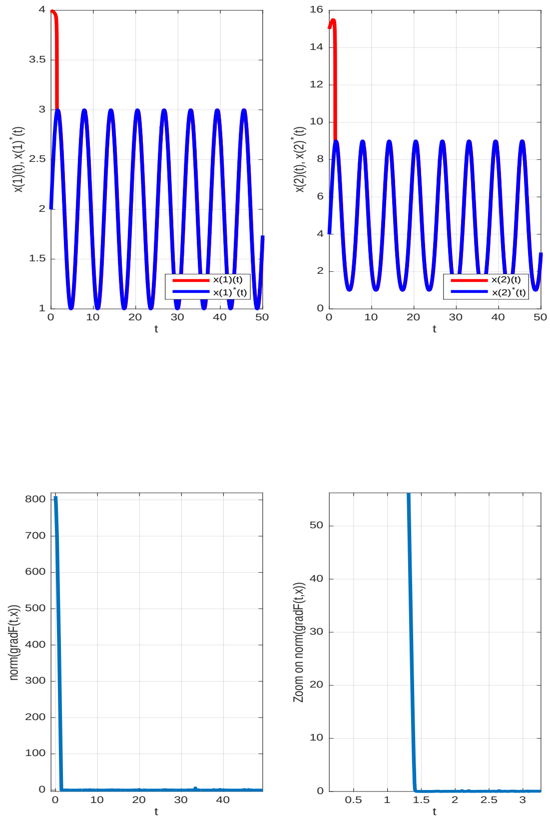

Fig. 6. Time-varying cost function: Flow (11) with $l(t,x) = 500$.

5. CONCLUSION

We have introduced a new family of second-order flows for continuous-time optimization of time-varying cost functions. The main characteristic of the proposed flows is their pre-defined finite-time convergence guarantees. To be able to analyze these discontinuous flows, we had to first extend an exiting Lyapunov-based inequality condition for finite-time stability in the case of smooth dynamics to the case of non-smooth dynamics modeled by time-varying differential inclusions. These flows were tested on two well known optimization benchmarks.

Future investigation will include working on developing constructive discretization methods, which preserve the main guarantees of the proposed flows.

APPENDIX

Recall that a function $x : I \to \mathbb{R}^n$ defined over an interval $I \subset \mathbb{R}$ is *absolutely continuous* if, for every $\varepsilon > 0$, there exists some $\delta > 0$ such that

$$\sum_{j=1}^{k}(t_j' - t_j) < \delta \implies \sum_{j=1}^{k}\|x(t_j') - x(t_j)\| < \varepsilon \tag{16}$$

for any disjoint subintervals $[t_1, t_1'], \ldots, [t_k, t_k'] \subseteq I$.

Lemma 1. If $x : I \to \mathbb{R}^n$ is absolutely continuous, then so is $t \mapsto (t, x(t))$.

Proof 3. We start by fixing an arbitrarily small $\varepsilon > 0$. Since $x(\cdot)$ is absolutely continuous, we can choose some $\delta > 0$ such that (16) holds. Furthermore, we can clearly always make δ smaller, and thus assume $0 < \delta \leq \varepsilon$. Let $\varepsilon' = \varepsilon - \delta$. Once again invoking the absolute continuity of $x(\cdot)$, we can choose some $\delta' > 0$ such that (16) holds for δ' and ε' instead of δ and ε. Furthermore, we can choose δ' in the interval $(0, \delta]$. Therefore, we have, for any disjoint subintervals $[t_1, t_1'], \ldots, [t_k, t_k'] \subset I$ such that $\sum_{j=1}^{k}(t_j' - t_j) < \delta$,

$$\sum_{j=1}^{k}\|(t_j', x(t_j')) - (t_j, x(t_j))\| \tag{17a}$$

$$= \sum_{j=1}^{k}\|(t_j' - t_j, x(t_j') - x(t_j))\| \tag{17b}$$

$$\leq \sum_{j=1}^{k}[(t_j' - t_j) + \|x(t_j') - x(t_j)\|] \tag{17c}$$

$$< \delta' + \varepsilon' \tag{17d}$$

$$\leq \varepsilon. \tag{17e}$$

Therefore, $t \mapsto (t, x(t))$ is absolutely continuous in I.

As a direct corollary, we have the following result.

Lemma 2. If $x : [0, \tau] \to \mathbb{R}^n$ is absolutely continuous and $V : [0, \tau] \times \mathcal{D} \to \mathbb{R}^n$ is Lipschitz continuous (in both variables), where $\mathcal{D} \subset \mathbb{R}^n$ is an open set that contains the trajectory $x(\cdot)$, then $t \mapsto V(t, x(t))$ is absolutely continuous.

Proof 4. By Lemma 1, we know that $t \mapsto (t, x(t))$ is absolutely continuous in $[0, \tau]$. Therefore, given that $V(\cdot)$ is Lipschitz continuous, it follows that its composition with $t \mapsto (t, x(t))$ is absolutely continuous in $[0, \tau]$.

REFERENCES

W. Su, S. Boyd, and E. J. Candes, "A differential equation for modeling Nesterovs accelerated gradient method: Theory and insights," *Journal of Machine Learning Research*, no. 17, pp. 1–43, 2016.

J. Cortés, "Finite-time convergent gradient flows with applications to network consensus," *Automatica*, vol. 42, pp. 1993–2000, 03 2006.

J. I. Poveda and N. Li, "Inducing Uniform Asymptotic Stability in Non-Autonomous Accelerated Optimization Dynamics via Hybrid Regularization', *Proceedings of the 58th IEEE Conference on Decision and Control* (to appear), 2019.

O. Romero and M. Benosman, "Finite-Time Convergence of Continuous-Time Optimization Algorithms via Differential Inclusions', *Neurips Workshop 'Beyond first order methods in machine learning systems'* (to appear), 2019.

G. Franca, D. Robinson, and R. Vidal, "A dynamical systems perspective on nonsmooth constrained optimization," *arXiv:1808.04048v2 [math.OC]*, 2019(a).

——, "ADMM and accelerated ADMM as continuous dynamical systems," *Proceedings of the 35th International Conference on Machine Learning, Stockholm*, 2018.

J. Poveda and A. Teel, "A framework for a class of hybrid extremum seeking controllers with dynamic inclusions," *Automatica*, no. 76, pp. 113–126, 2017.

M. Guay and T. Zhang, "Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties," *Automatica*, vol. 39, pp. 1283–1293, 2003.

M. Krstić, "Stability of extremum seeking feedback for general nonlinear dynamic agents," *Automatica*, vol. 36, pp. 595–601, 2000.

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

A. Scheinker and M. Krstić, *Model-free stabilization by extremum seeking*. Springer, 2016.

C. Zhang and R. Ordónez, *Extremum-Seeking Control and Applications*. Springer, 2012.

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," *Automatica*, no. 94, pp. 151–160, 2018.

M. Benosman and K.-Y. Lum, "Nonlinear control allocation for non-minimum phase systems," *IEEE Transactions on Control Systems Technology*, vol. 17, no. 2, pp. 394–404, 2009.

B. Paden and S. Sastry, "A calculus for computing Filippov as differential inclusion with application to the variable structure control of robot manipulators, " *Circuits and Systems, IEEE Transactions on*, vol. 34, pp. 73 – 82, 02 1987.

F. Clarke, "Nonsmooth analysis in control theory: A survey," *European Journal of Control*, vol. 7, no. 2, pp. 145 – 159, 2001.

A. Bacciotti and F. Ceragioli, "Stability and stabilization of discontinuous systems and nonsmooth lyapunov functions," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 4, pp. 361–376, 1999.

A. F. Filippov and F. M. Arscott, "Differential equations with discontinuous righthand sides," in *Mathematics and its Applications*, Springer, 1988.

J. Cortes, "Discontinuous dynamical systems," *IEEE Control Systems Magazine*, vol. 28, no. 3, pp. 36–73, June 2008.

Y. Yuan, M. Li, J. Liu, and C. Tomlin,, "On the Powerball Method: Variants of Descent Methods for Accelerated Optimization," *IEEE Control Systems Letters*, vol. 3, no. 3, pp. 601–606, July 2019.