Continuous-Time Optimization of Time-Varying Cost Functions via Finite-Time Stability with Pre-Defined Convergence Time

Romero, Orlando; Benosman, Mouhacine

TR2020-085 July 01, 2020

Abstract

In this paper, we propose a new family of continuous-time optimization algorithms for time-varying, locally strongly convex cost functions, based on discontinuous second-order gradient optimization flows with provable finite-time convergence to local optima. To analyze our flows, we first extend a well-know Lyapunov inequality condition for finite-time stability, to the case of arbitrary time-varying differential inclusions, particularly of the Filippov type. We then prove the convergence of our proposed flows in finite time. We illustrate the performance of our proposed flows on a quadratic cost function to track a decaying sinusoid.

American Control Conference (ACC)

Continuous-Time Optimization of Time-Varying Cost Functions via Finite-Time Stability with Pre-Defined Convergence Time

Orlando Romero[†] Mouhacine Benosman[‡]

***Abstract*— In this paper, we propose a new family of continuous-time optimization algorithms for time-varying, locally strongly convex cost functions, based on discontinuous second-order gradient optimization flows with provable *finite-time* convergence to local optima. To analyze our flows, we first extend a well-know Lyapunov inequality condition for finite-time stability, to the case of arbitrary time-varying differential inclusions, particularly of the Filippov type. We then prove the convergence of our proposed flows in finite time. We illustrate the performance of our proposed flows on a quadratic cost function to track a decaying sinusoid.**

I. Introduction

In continuous-time optimization, an ordinary differential equation (ODE) or a partial differential equation (PDE) is designed in such a way that its solution convergences over time to an optimal value of the cost function. There has been a recent surge in research papers in this direction, arguably starting with the pioneer work by Brockett in [1], such as [2]–[4], [6], [7], [9]–[23]

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, typically modeled as the measurable 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., [20]–[23]. 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 [4], 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 Nesterov's algorithm. More recently, in [5], the authors establish uniform asymptotic stability and robustness properties for the continuous-time limit of the Nesterovs accelerated gradient method, by using resetting mechanisms that are modeled by well-posed hybrid dynamical systems.

In [19], 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 [18], 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 [17], the authors study the crucial problem of structure-preserving discretizations of continuous-time optimization flows. More specifically, the authors focus on two classes of conformal Hamiltonian systems whose trajectories lie on a symplectic manifold, namely a classical mechanical system with linear dissipation and its relativistic extension. One of the most noticeable claims in this paper is that conformal symplectic integrators can preserve convergence rates of the continuous-time system up to a negligible error. As a by product of this, the authors show that the classical momentum method is a symplectic integrator. Finally, a relativistic generalization of classical momentum called *relativistic gradient descent* is introduced, and it is argued that it may result in more stable/faster optimization for some optimization problems.

In [7], 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. In [8] the authors proposed a class of discontinuous dynamical systems, of second order with respect to the cost function, that are continuous-time optimization algorithms with finite-time convergence and prescribed convergence time.

In this work, we want to focus on this specific class

[†]Department of Industrial and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA, rodrio2@rpi.edu.
[‡]Mitsubishi Electrics Research Laboratories, Cambridge, MA 02139, USA, benosman@merl.com.

of continuous-time optimization algorithms with finite-time convergence. We propose a new family of *discontinuous* second-order flows, which guarantee local convergence to an optimum trajectory of a time-varying cost function, in a *pre-defined, finite time.*

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 [24], 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 differential inclusions.

This paper is organized as follows: Section II is dedicated to some preliminaries about continuous-time optimization, and finite-time stability in the context of differential inclusions. Our main results are presented in Section III, 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 IV, we show the efficiency of this continuous-time optimization flow with a numerical example. The paper ends with a summarizing conclusion and a discussion of our ongoing investigations, in Section V.

Notation

We use the notation $\triangleq$ to defined functions or other mathematical objects. $\mathbb{R}$ and $\mathbb{R}_+$ denote, respectively, the set of real numbers and set of non-negative real numbers. Given a time-varying cost function $f(t,x) : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$, we denote its gradient and Hessian matrix with respect to x as $\nabla f(t,x)$ and $\nabla^2 f(t,x)$, consecutively. Given two sets X and Y, the notation $K : X \rightrightarrows Y$ denotes a set-valued map K such that $K(x) \subseteq Y$ for each $x \in X$.

II. Preliminaries and Problem Statement

In this section, we will review some key ideas regarding discontinuous state-space dynamical systems with a focus on Filippov differential inclusions. Then, we will review key ideas and definitions regarding finite-time stability of time-varying differential inclusions with respect to Carathéodory solutions. Finally, we will formulate and formally state the problem addressed in this paper.

A. Discontinuous Systems; Filippov Differential Inclusions

Recall that 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 [27], [31] 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 (*i.e.* closure of the convex hull). 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 [27], [31]. Furthermore, the Filippov set-valued map (3) can be computed as

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

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, and where co denotes the convex hull. 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 $K[-\nabla f](t,x) = \{-\nabla f(x)\}$ for every $x \in \mathbb{R}^n$, provided that f is continuously differentiable. Furthermore, if f is only Lipschitz continuous, then $K[-\nabla f](t,x) = -\partial f(x)$, where ∂f denotes Clarke's generalized gradient [28].

B. Finite-Time Stability for Differential Inclusions

Consider a general differential inclusion [29]

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

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

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 [30] 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 $t\to T(x_0)$

C. Problem Statement

Consider some time-varying objective cost function $f:\mathbb{R}_+\times\mathbb{R}^n\to\mathbb{R}$ that we wish to minimize. In particular, let $x^\star(t)\in\mathbb{R}^n$ be a local minimum of $f(t,\cdot)$ that is unknown to us but under the assumption that $x^\star(\cdot)$ is sufficiently smooth. In continuous-time optimization, we typically proceed by designing a nonlinear state-space dynamical system

$$\dot{x}=F(t,x) \tag{6}$$

for which $F(t,x)$ can be computed without knowledge of $x^\star$ and with limited knowledge or access to f (*e.g.* using only current and local information on f, up to second order). Traditionally, the system (6) is often implicitly or explicitly designed so as to make

$$\dot{e}=F_{\text{error}}(t,e), \tag{7}$$

with

$$F_{\text{error}}(t,e)\triangleq F(t,x^\star(t)+e)-\dot{x}^\star(t), \tag{8}$$

asymptotically stable at the origin. Naturally, $e(t)\triangleq x(t)-x^\star(t)$ simply amounts to a tracking error such that systems‘(6) and (7) are equivalent.

Substantial work has been done for this purpose, with much of it focused in optimizing the *asymptotic* convergence rate for $x(t)\to x^\star$ as $t\to\infty$ and its translation into a discretized scheme suitable for modern digital computing. In this work, however, we seek continuous-time dynamical systems that are possibly discontinuous or non-Lipschitz (and thus based on differential inclusions as opposed to exclusively ODEs), for which (7) is *finite-time* stable at the origin. Furthermore, we seek design systems using only up to second-order information on f, and in such a way that the finite-time *settling time* can be prescribed, or at least controlled to some degree with the aforementioned limited information.

Our approach to achieve this objective[1] is largely based on exploiting the Lyapunov-like differential inequality

$$\dot{\mathcal{E}}(t)\leq -c\,\mathcal{E}(t)^\alpha,\quad \text{a.e. } t\geq 0, \tag{9}$$

with constants $c>0$ and $\alpha<1$, for absolutely continuous functions $\mathcal{E}(\cdot)$ such that $\mathcal{E}(0)>0$. Indeed, under the aforementioned conditions, $\mathcal{E}(t)$ will reach $\mathcal{E}(t)=0$ in finite time $t=t^\star\leq\frac{\mathcal{E}(0)^{1-\alpha}}{c(1-\alpha)}<\infty$, with this upper-bound being an equality whenever (9) is an equality as well (see e.g. Lemma 1 in [24]).

First, we must propose a sufficiently smooth candidate Lyapunov function $V=V(t,e)$, which must be (locally) positive definite. In other words, $(t,x)\mapsto V(t,x-x^\star(t))$ can potentially serve as a convenient surrogate for the objective function f. Candidate Lyapunov functions include

$$V(t,x-x^\star(t))=\|x-x^\star(t)\|^2 \iff V(t,e)=\|e\|^2, \tag{10}$$

$$V(t,x-x^\star(t))=f(t,x)-f(t,x^\star(t)) \tag{11a}$$

$$\iff V(t,e)=f(t,x^\star(t)+e)-f(t,x^\star(t)), \tag{11b}$$

and

$$V(t,x-x^\star(t))=\|\nabla f(t,x)\|^2 \tag{12a}$$

$$\iff V(t,e)=\|\nabla f(t,x^\star(t)+e)\|^2 \tag{12b}$$

with (11b) requiring $x^\star(t)$ to be a strict local minimum of $f(t,\cdot)$ for each $t\geq 0$, and (12b) requiring $x^\star(t)$ to be an isolated stationary point of $f(t,\cdot)$ for each $t\geq 0$ (*e.g.* if $f(t,\cdot)$ is strongly convex near $x^\star(t)$). Then, the objective is to design a (possibly discontinuous) function F that can be computed without explicit knowledge of $x^\star(\cdot)$, and using only up to second-order information on f such that, for every Filippov solution $x(\cdot)$ of (6), the inequality (9) is satisfied for the energy function $\mathcal{E}(t)\triangleq V(t,x(t)-x^\star(t))$. More precisely, we seek F such that (7) is finite-time stable at the origin.

We now summarize the problem statement.

Problem 1. *Given a sufficiently smooth time-varying cost function $f:\mathbb{R}_+\times\mathbb{R}^n\to\mathbb{R}$ with a sufficiently regular local time-varying minimizer $x^\star(\cdot)$, solve the following tasks:*

1) *Design a sufficiently smooth[2] candidate Lyapunov function $V(t,e)$ such that $(t,x)\mapsto V(t,x-x^\star(t))$ is surrogate to f near $(t,x^\star(t))$;*
2) *Design a (possibly discontinuous) system[3] (6) such that $F(t,x)$ can be computed using only the information*

[1] Other approaches to establish finite-time stability for discontinuous systems based on the notion of homogeneity have been presented, e.g., in [33] and references therein.

[2] At least locally Lipschitz continuous and regular, but continuous differentiability suffices.

[3] Right-hand side (RHS) defined at least a.e., Lebesgue measurable, and locally essentially bounded.

available on f and such that (8) *is finite-time stable at the origin.*

By following this strategy, we will therefore achieve (local) *finite-time convergence*. Furthermore, if $V(0, x_0 - x^\star(0))$ computed or upper bounded, then F can be readily tuned to achieve finite-time convergence under a *prescribed* range for the settling time, or with *exact* prescribed settling time if $V(0, x_0 - x^\star(0))$ can be explicitly computed and (9) holds exactly for $\mathcal{E}(t) \triangleq V(t, x(t) - x^\star(t))$ and any Filippov solution $x(\cdot)$.

III. Main results

In this section, we start by establishing a Lyapunov-based sufficient condition for the finite-time stability of general time-varying differential inclusions. Then, given a time-varying cost function, we construct a family of second-order flows such that the system associated with the tracking error with respect to the time-varying local minimum is finite-time stable at the origin.

A. A Sufficient Lyapunov Condition for Finite-Time Stability of Differential Inclusions

We will now construct a Lyapunov-based criterion adapted from the literature of finite-time stability of Lipschitz continuous systems into general differential inclusions. This results is based on Lemma 1 in [24], which can be adapted to include absolutely continuous solutions and non-positive exponents. We refer also to (Proposition 5, [32]) for a similar result for the case of time-invariant systems.

Theorem 1. *Let $K : \mathbb{R}_+ \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued map satisfying Assumption 2 and $V : \mathbb{R}_+ \times \mathcal{D} \to \mathbb{R}$ a continuously differentiable function such that $V(t, \cdot)$ is positive definite for every $t \geq 0$, where $\mathcal{D} \subset \mathbb{R}^n$ is an open and positively invariant neighborhood of the origin. Let*

$$\dot{V}(t,x) \triangleq \left\{ \frac{\partial V}{\partial t}(t,x) + \nabla V(t,x) \cdot v : v \in K(t,x) \right\} \quad (13)$$

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 $\alpha < 1$ such that

$$\max \dot{V}(t,x) \leq -c\, V(t,x)^\alpha \quad (14)$$

a.e. in $t \geq 0$ and $x \in \mathcal{D}$, then (5) *is finite-time stable at the origin with settling time upper bounded by*

$$T(x_0) \leq \frac{V(0,x_0)^{1-\alpha}}{c(1-\alpha)} \quad (15)$$

for each $x_0 \in \mathcal{D}$. Furthermore, if $\dot{V}(t,x)$ is a singleton a.e. in $t \geq 0$ and $x \in \mathcal{D}$, and (14) *is an exact equality, then so is* (15).

Proof. Lyapunov stability follows from [30, 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 due to V being continuously differentiable. Therefore, since $\frac{d}{dt}V(t,x(t)) = \dot{\mathcal{E}}(t) \in \dot{V}(t, x(t))$ [29, Lemma 1], we note from (14) that

$$\frac{d}{dt}V(t,x(t)) \leq -cV(t,x(t))^\alpha, \quad (16)$$

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

B. Proposed Family of Finite-Time Converging Flows

We are ready to propose some optimization flows with finite-time convergence guarantees. The first family of flows is in the form of Newton-like discontinuous flows with *pre-defined finite* settling time $T > 0$.

First, let us state a smoothness assumption on the cost function and time-varying minimizer of interest.

Assumption 3. $f : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable, and there exist some $m > 0$ and some continuously differentiable $x^\star(\cdot)$ such that $x^\star(t)$ is a stationary point of $f(t, \cdot)$ and $f(t, \cdot)$ is m-strongly convex near $x^\star(t)$, for each $t \geq 0$.

We are now ready to state our proposed family of continuous-time optimization algorithms, and thus our main result.

Theorem 2. *Let $T > 0$, $\alpha \in [1/2, 1)$, and $r \in \mathbb{R}$. Under Assumption 3, any maximal Filippov solution $x(\cdot)$ of*

$$\dot{x} = -w(t,x;x_0)\frac{[\nabla^2 f(t,x)]^r \nabla f(t,x)}{\nabla f(t,x)^\top [\nabla^2 f(t,x)]^{r+1} \nabla f(t,x)}, \quad (17)$$

with

$$w(t,x;x_0) \triangleq \frac{\|\nabla f(0,x_0)\|^{2(1-\alpha)}}{2(1-\alpha)T} \|\nabla f(t,x)\|^{2\alpha} + \nabla f(t,x)^\top \frac{\partial \nabla f}{\partial t}(t,x) \quad (18)$$

and $x(0) = x_0$ sufficiently close to $x^\star(0)$ will converge in finite time to $x^\star$, with an exact settling time of $T(x_0) = T$.

Proof. Consider the system (7), where $F(t,x)$ denotes the RHS of (17). Clearly, given any Filippov solution $e(\cdot)$ to (7), then $x(t) \triangleq e(t) + x^\star(t)$ is likewise a Filippov solution to (17). The strategy that follows now, is to stablish finite-time stability of (7) by leveraging Theorem 1.

Let $\mathcal{D}$ be an open and positively invariant set (w.r.t. (17)) that contains $\{x^\star(t) : t \geq 0\}$, over which $x^\star(t)$ is the only stationary point of $f(t, \cdot)$ and $f(t, \cdot)$ is m-strongly convex, for each $t \geq 0$ and $x \in \mathcal{D}$. Let $\mathcal{D}' = \{x - x^\star(0) : x \in \mathcal{D}\}$ and $V : \mathbb{R}_+ \times \mathcal{D}' \to \mathbb{R}$ be given by $V(t,e) \triangleq \|\nabla f(t, x^\star(t) + e)\|^2$. Clearly, $\mathcal{D}'$ is positively invariant w.r.t. to (7), and V is continuously differentiable and positive definite.

Let us now check that F, and subsequently F_{error}, both satisfy Assumption 1. Indeed, notice that $F(t,x)$ is continuous everywhere in $\mathbb{R}_+ \times \mathcal{D}$ except for $x = x^\star(t)$, where it is possibly undefined. Therefore, F is measurable. On the other hand, given $(t,x) \in \mathbb{R}_+ \times \mathcal{D}$ such that $x \neq x^\star(t)$, we have

$$\|F(t,x)\| = \frac{|w(t,x;x_0)| \| [\nabla^2 f(t,x)]^r \nabla f(t,x) \|}{\nabla f(t,x)^\top [\nabla^2 f(t,x)]^{r+1} \nabla f(t,x)} \tag{19a}$$

$$\leq \frac{|w(t,x;x_0)| \lambda_{\max}[\nabla^2 f(t,x)]^r \|\nabla f(t,x)\|}{\lambda_{\min}[\nabla^2 f(t,x)]^{r+1} \|\nabla f(t,x)\|^2} \tag{19b}$$

$$= \left(\frac{|w(t,x;x_0)|}{\|\nabla f(t,x)\|} \right) \frac{\lambda_{\max}[\nabla^2 f(t,x)]^r}{\lambda_{\min}[\nabla^2 f(t,x)]^{r+1}}, \tag{19c}$$

with

$$\frac{|w(t,x;x_0)|}{\|\nabla f(t,x)\|} \leq \frac{\|\nabla f(0,x_0)\|^{2(1-\alpha)}}{2(1-\alpha)T} \|\nabla f(t,x)\|^{2\alpha - 1} + \left\| \frac{\partial \nabla f}{\partial t}(t,x) \right\|. \tag{20}$$

Now, since f is twice continuously differentiable and $\alpha \geq 1/2 \iff 2\alpha - 1 \geq 0$, it follows that the RHS of (20) is continuous. Furthermore, since $f(t,\cdot)$ is m-strongly convex, and thus $\lambda_{\min}[\nabla^2 f(t,x)] \geq m > 0$ in $\mathcal{D}$, then the RHS of (19c) is continuous as well. Putting everything together, we see that $\|F\|$ is upper bounded by a continuous function, and thus F must be locally essentially bounded. Therefore, Assumption 1 is indeed satisfied for F, and thus F_{error} as well.

Proceeding, we now fix some arbitrary $t \geq 0$ and $e = x - x^\star(t) \in \mathcal{D}' \setminus \{0\}$. We thus have

$$\frac{\partial V}{\partial t}(t,e) = 2\nabla f(t,x)^\top \left[\frac{\partial \nabla f}{\partial t}(t,x) + \nabla^2 f(t,x) \dot{x}^\star(t) \right]; \tag{21}$$

$$\nabla V(t,e) = 2\nabla^2 f(t,x) \nabla f(t,x); \tag{22}$$

$$K(t,e) \triangleq \mathcal{K}[F_{\text{error}}](t,e) = \{F(t,x) - \dot{x}^\star(t)\}. \tag{23}$$

Therefore,

$$\dot{V}(t,e) = \left\{ \frac{\partial V}{\partial t}(t,e) + \nabla V(t,e) \cdot v : v \in K(t,e) \right\} \tag{24}$$

is a singleton satisfying

$$\max \dot{V}(t,e) = 2\nabla f(t,x)^\top \left[\frac{\partial \nabla f}{\partial t}(t,x) + \nabla^2 f(t,x) \dot{x}^\star(t) \right] + 2\nabla f(t,x)^\top \nabla^2 f(t,x) (F(t,x) - \dot{x}^\star(t)) \tag{25a}$$

$$= 2\nabla f(t,x)^\top \left[\frac{\partial \nabla f}{\partial t}(t,x) + \nabla^2 f(t,x) F(t,x) \right] \tag{25b}$$

$$= 2\nabla f(t,x)^\top \frac{\partial \nabla f}{\partial t}(t,x) - 2w(t,x;x_0) \tag{25c}$$

$$= \underbrace{\left(\frac{\|\nabla f(0,x_0)\|^{2(1-\alpha)}}{(1-\alpha)T} \right)}_{\triangleq c(x_0)} \underbrace{\|\nabla f(t,x)\|^{2\alpha}}_{= V(t,e)^\alpha}, \tag{25d}$$

with $c(x_0) > 0$ independent of (t,e). The result follows by invoking Theorem 1. ■

Remark 1 (Convergence Domain)**.** The convergence of Filippov solutions $x(\cdot)$ of (17) in finite time holds true for any x_0 in the domain $\mathcal{D}$ constructed in the previous proof.

Remark 2 (Range of α)**.** In Theorem 2, we selected the exponent parameter $\alpha \in [1/2, 1)$, thus discarding $\alpha < 1/2$, in order to ensure F to be locally essentially bounded and thus ensure that Filippov solutions to our proposed family of flows exist. Also notice that for $\alpha \in (1/2, 1)$, traditional solutions exist since F is continuous for that range of α. However, $F(t,\cdot)$ remains non-Lipschitz and thus uniqueness of solutions is not immediately guaranteed. For $\alpha = 1/2$, however, only Filippov solutions appear to be meaningful.

Remark 3 (Practical Stability)**.** In a practical implementation, if we want to keep all the remaining range of α, i.e., $\alpha < 0.5$, we could simply incorporate a *regularization term* in the flow (17), as

$$\dot{x} = -\frac{w(t,x;x_0)[\nabla^2 f(t,x)]^r \nabla f(t,x)}{\delta + \nabla f(t,x)^\top [\nabla^2 f(t,x)]^{r+1} \nabla f(t,x)}, \tag{26}$$

where $\delta > 0$ is a small constant. From a theoretical convergence perspective, this implementation 'fix' will simply change the finite-time convergence, to a finite-time *practical* convergence, i.e., given any arbitrarily small $\varepsilon > 0$, we can choose $\delta > 0$ sufficiently small for the tracking error to reach an ε-neighborhood of zero, and thus 'practically' vanish. In practical terms, such a regularized flow is more suitable for numerical solvers, so more work will be conducted into studying, from a numerical optimization perspective, fast discretization schemes for (26).

IV. Numerical Examples

We will now test our proposed flows on a simple time-varying cost function – a convex quadratic cost function that to track a decaying sinuisoid. More precisely, let us consider the time-varying cost function

$$f(t,x) \triangleq (x - x^\star(t))^2, \tag{27}$$

with $x^\star(t) \triangleq 10 \sin\left(\frac{4t}{1+0.1t^2}\right)$, which together satisfy Assumption 3.

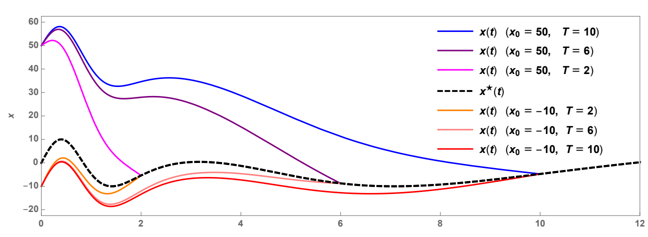

Fig. 1: Trajectories of the proposed flow (17) with parameters $(\alpha, r) = (1/2, -1)$ for (27), with time-varying minimum $x^\star(\cdot)$. Different initial approximations and pre-defined settling times are illustrated.

As we can see in Figure 1, our proposed method is able to naturally deal with time-varying cost functions. Furthermore, we can see that the tracking error reaches zero exactly at the pre-defined time $T > 0$, regardless of the initial value $x_0 \in \mathbb{R}$, since $f(t,\cdot)$ is strongly convex everywhere, for every $t \geq 0$.

V. CONCLUSION AND FUTURE WORK

We have introduced a new family of discontinuous second-order flows for continuous-time optimization. The main characteristic of the proposed flows is their finite-time convergence guarantees, with an arbitrary *pre-defined* (by the user) convergence time. To analyze these discontinuous flows, we first extended 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. We then derived and established finite-time stability (and thus convergence) for the proposed family of continuous-time optimization algorithms. Finally, we conducted numerical experiments on a time-varying quadratic cost function that tracks a decaying sinusoid.

While the obtained results are encouraging, there are several unanswered questions, which we will target in our future work. First, while we have used commonly available numerical solvers in our (small-scale) numerical experiments, it is evidently clear that traditional ODE solvers do not appear to translate into competitive *iterative* optimization algorithms. More work will be done in this discretization research direction. Furthermore, we also seek to adapt our methods to be based on first-order information on the cost function, as well as to a allow for linear and nonlinear constraints, and to develop distributed and decentralized variants. Lastly, many real-life problems that require a time-varying optimization framework, such as in motion planning or formation control in robotics, do not allow direct access to gradients, Hessian matrices, or time-derivatives of the gradient. Instead, these are typically estimated based on measurements (e.g. of the cost function) that often occur in discrete time and carry noisy perturbations. Therefore, future work will also be dedicated to the robustification of our proposed flows, including zeroth-order (gradient-free) schemes, for example, based on continuous-time finite-difference schemes, or on an extremum-seeking formulation of the present flows.

REFERENCES

[1] R. Brockett, "Dynamical systems that sort lists, diagonalize matrices and solve linear programming problems," in *IEEE Conference on Decision and Control*, 1988, pp. 799–803.

[2] D. Scieur, V. Roulet, F. Bach, and A. d'Aspremont, "Integration methods and accelerated optimization algorithms," *31st Conference on Neural Information Processing Systems*, 2017.

[3] J. Zhang, A. Mokhtari, S. Sra, and A. Jadbabaie, "Direct runge-kutta discretization achieves acceleration," in *32st Conference on Neural Information Processing Systems*, 2018.

[4] 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.

[5] Jorge I. Poveda and Na 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.

[6] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," in *IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 3800–3805.

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

[8] Romero Orlando, Benosman Mouhacine. Finite-Time Convergence of Continuous-Time Optimization Algorithms via Differential Inclusions. In: *the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), workshop: Beyond first order methods in machine learning systems*:1–11; 2019.

[9] C. Wang and Y. M. Lu, "The scaling limit of high-dimensional online independent component analysis," in *31st Conference on Neural Information Processing Systems*, 2017.

[10] A. C. Wilson, B. Recht, and M. I. Jordan, "A lyapunov analysis of momentum methods in optimization," *arXiv:1611.02635*, 2016.

[11] J. Wang and N. Elia, "Control approach to distributed optimization," in *48th Annual Allerton Conference*, 2010, pp. 557–561.

[12] L. Faybusovich, "Dynamical systems which solve optimization problems with linear constraints," *IMA Journal of Mathematical Control & Information*, pp. 135–149, 1991.

[13] H. Attouch and J. Peypouquet, "The rate of convergence of nesterov's accelerated forward-backward method is actually $o(k^{-2})$," *SIAM Journal on Optimization*, no. 3, pp. 1–7, 2015.

[14] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*, 1996.

[15] H. Attouch, J. Peypouquet, and P. Redont, "Fast convex optimization via inertial dynamics with hessian driven damping," *Journal of Differential Equations*, no. 10, pp. 5734–5783, 2016.

[16] ——, "Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity," *Mathematical Programming*, no. 168, pp. 123–175, 2018.

[17] G. Franca, J. Sulam, D. Robinson, and R. Vidal, "Conformal symplectic and relativistic optimization," *arXiv:1903.04100v3 [math.OC]*, 2019.

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

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

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

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

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

[23] 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.

[24] 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.

[25] Y.-J. Liu and S. Tong, "Barrier Lyapunov functions-based adaptive control for a class of nonlinear pure-feedback systems with full state constraints," *Automatica*, no. 64, pp. 70–75, 2016.

[26] Y. Yang, K. G. Vamvoudakis, H. Modares, W. He, Y. Yin, and D. C. Wunsch, "Safety-aware reinforcement learning framework with an actor-critic-barrier structure," in *IEEE American Control Conference*, 2019.

[27] 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.

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

[29] 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.

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

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

[32] E. Moulay and W. Perruquetti, "Finite time stability of differential inclusions," *IMA Journal of Mathematical Control and Information*, vol. 22, pp. 465–475, June 2005.

[33] A. Polyakov, D. Efimov and B. Brogliato, "Consistent discretization of finite-time and fixed-time stable systems," *SIAM Journal on Control and Optimization*, vol. 57, no. 1, pp. 78–103, 2019.