Finite-time convergent gradient flows with applications to network consensus ⋆

Jorge Cortés

Department of Applied Mathematics and Statistics, Baskin School of Engineering, University of California, Santa Cruz, CA 95064, USA, email: `jcortes@ucsc.edu`

Abstract

This paper introduces the normalized and signed gradient dynamical systems associated with a differentiable function. Extending recent results on nonsmooth stability analysis, we characterize their asymptotic convergence properties and identify conditions that guarantee finite-time convergence. We discuss the application of the results to consensus problems in multi-agent systems and show how the proposed nonsmooth gradient flows achieve consensus in finite time.

Key words: Gradient flows; Nonsmooth analysis; Finite-time convergence; Network consensus; Multi-agent systems

1 Introduction

Consider the gradient flow $\dot{x} = -\operatorname{grad}(f)(x)$ of a differentiable function $f : \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$. It is well known (see e.g. [13]) that the minima of f are stable equilibria of this system, and that, if the level sets of f are bounded, then the trajectories converge asymptotically to the set of critical points of f. Gradient dynamical systems are employed in a wide range of applications, including optimization, parallel computing and motion planning. In robotics, potential field methods are used to autonomously navigate a robot in a cluttered environment. Gradient algorithms enjoy many features: they are naturally robust to perturbations and measurement errors, amenable to asynchronous implementations, and admit efficient numerical approximations.

In this note, we provide an answer to the following question: how could one modify the gradient vector field above so that the trajectories converge to the critical points of the function *in finite time*? - as opposed to over an infinite-time horizon. There are a number of settings where finite-time convergence is a desirable property. We study this problem with the aim of designing gradient coordination algorithms for multi-agent systems that achieve the desired task in finite time.

Our answer to the question above is the flows

$$\dot{x} = -\frac{\operatorname{grad}(f)(x)}{\|\operatorname{grad}(f)(x)\|_2}, \qquad \dot{x} = -\operatorname{sgn}(\operatorname{grad}(f)(x)),$$

where $\|\cdot\|_2$ denotes the Euclidean distance and $\operatorname{sgn}(x) = (\operatorname{sgn}(x_1), \dots, \operatorname{sgn}(x_d))$. Using tools from nonsmooth stability analysis, we show that, under some assumptions on f, both systems are guaranteed to achieve the set of critical points in finite time.

Literature review. Guidelines on how to design dynamical systems for optimization purposes, with a special emphasis on gradient systems, are described in [12]. The book [2] thoroughly discusses gradient descent flows in distributed computation in settings with fixed-communication topologies. Nonsmooth analysis studies the notion and computational properties of the generalized gradient [4]. Tools for establishing stability and convergence properties of nonsmooth dynamical systems via Lyapunov functions are presented in [1, 5, 6, 19] (see also references therein). Finite-time discontinuous feedback stabilizers for a class of planar systems are proposed in [18]. Finite-time stability of continuous autonomous systems is rigorously studied in [3]. The reference [7] develops finite-time stabilization strategies based on time-varying feedback. Previous work on motion coordination of multi-agent systems has proposed cooperative algorithms based on gradient flows to achieve tasks such as cohesiveness [11, 14, 20], deployment [8, 9] and consensus [15, 16, 17], The distributed algorithms in these works achieve the desired coordination task over an infinite-time horizon.

Statement of contributions. In this paper we introduce the normalized and signed gradient descent flows associated to a differentiable function. We characterize their convergence properties via nonsmooth stability analysis. We also identify general conditions under which

⋆ Research supported by NSF CAREER award ECS-0546871. This work was not presented at any IFAC meeting.

these flows reach the set of critical points of the function in finite time. To do this, we extend recent results on the convergence properties of general nonsmooth dynamical systems via locally Lipschitz and regular Lyapunov functions (e.g. [1, 8, 19]). In particular, we develop two novel results involving second-order information on the evolution of the Lyapunov function to establish finite-time convergence. These results are not restricted to gradient flows, and can indeed be used in other setups with discontinuous vector fields and locally Lipschitz functions. We discuss in detail the application of these results to network consensus problems. We propose two coordination algorithms based on the Laplacian of the network graph that achieve consensus in finite time. The normalized gradient descent of the Laplacian potential is not distributed over the network graph and achieves average-consensus, i.e., consensus on the average of the initial agents' states. The signed gradient descent of the Laplacian potential is distributed over the network graph and achieves average-max-min-consensus, i.e., consensus on the average of the maximum and the minimum values of the initial agents' states. We also consider networks with switching connected network topologies.

Organization. Section 2 introduces differential equations with discontinuous right-hand sides and presents various nonsmooth tools for stability analysis. Section 3 introduces the normalized and signed versions of the gradient descent flow of a function and characterizes their convergence properties. Conditions are given under which these flows converge in finite time. Section 4 discusses the application of the results to network consensus problems. Section 5 presents our conclusions.

Notation. The set of strictly positive natural (resp. real) numbers is denoted by $\mathbb{N}$ (resp. $\mathbb{R}_+$). For $d \in \mathbb{N}$, $e_1, \dots, e_d$ is the standard orthonormal basis of $\mathbb{R}^d$. For $x \in \mathbb{R}^d$, let $\operatorname{sgn}(x) = (\operatorname{sgn}(x_1), \dots, \operatorname{sgn}(x_d)) \in \mathbb{R}^d$, let x' be the transpose of x, and let $\|x\|_1$ and $\|x\|_2$ be the 1-norm and the 2-norm of x, resp. For $x, y \in \mathbb{R}^d$, let $x \cdot y$ be the inner product. Let $\mathbf{1} = (1, \dots, 1)' \in \mathbb{R}^d$. For $S \in \mathbb{R}^d$, let $\operatorname{co}(S)$ denote its convex closure. Define also $\operatorname{diag}((\mathbb{R}^d)^n) = \{(x, \dots, x) \in (\mathbb{R}^d)^n \mid x \in \mathbb{R}^d\}$ for $n \in \mathbb{N}$. Given a positive semidefinite $d \times d$-matrix A, let $H_0(A) \subset \mathbb{R}^d$ denote the eigenspace corresponding to 0 (if A is positive definite, set $H_0(A) = \{0\}$). We denote by $\pi_{H_0(A)} : \mathbb{R}^d \to H_0(A)$ the orthogonal projection onto $H_0(A)$. Let $\lambda_2(A)$ and $\lambda_d(A)$ be the smallest non-zero and greatest eigenvalue of A, resp., i.e. $\lambda_2(A) = \min\{\lambda \mid \lambda > 0 \text{ and } \lambda \text{ eigenvalue of } A\}$ and $\lambda_d(A) = \max\{\lambda \mid \lambda \text{ eigenvalue of } A\}$. It is easy to see

$$x' A\, x \geq \lambda_2(A)\, \|x - \pi_{H_0(A)}(x)\|_2^2, \quad x \in \mathbb{R}^d. \qquad (1)$$

2 Nonsmooth stability analysis

This section introduces discontinuous differential equations and presents various nonsmooth tools to analyze their stability properties. We present two novel results on the second-order evolution of locally Lipschitz functions and on finite-time convergence.

2.1 Discontinuous differential equations

For differential equations with discontinuous right-hand sides we understand the solutions in terms of differential inclusions following [10]. Let $F : \mathbb{R}^d \to 2^{\mathbb{R}^d}$ be a set-valued map. Consider the differential inclusion

$$\dot{x} \in F(x)\,. \qquad (2)$$

A solution to this equation on an interval $[t_0, t_1] \subset \mathbb{R}$ is defined as an absolutely continuous function $x : [t_0, t_1] \to \mathbb{R}^d$ such that $\dot{x}(t) \in F(x(t))$ for almost all $t \in [t_0, t_1]$. Now, consider the differential equation

$$\dot{x}(t) = X(x(t))\,, \qquad (3)$$

where $X : \mathbb{R}^d \to \mathbb{R}^d$ is measurable and essentially locally bounded [10]. We understand the solution of (3) in the Filippov sense. For each $x \in \mathbb{R}^d$, consider the set

$$K[X](x) = \bigcap_{\delta > 0} \bigcap_{\mu(S)=0} \operatorname{co}\{X(B_d(x, \delta) \setminus S)\}\,, \qquad (4)$$

where μ denotes the usual Lebesgue measure in $\mathbb{R}^d$. Intuitively, this set is the convexification of the limits of the values of the vector field X around x. A Filippov solution of (3) on an interval $[t_0, t_1] \subset \mathbb{R}$ is defined as a solution of the differential inclusion

$$\dot{x} \in K[X](x)\,. \qquad (5)$$

Since the set-valued map $K[X] : \mathbb{R}^d \to 2^{\mathbb{R}^d}$ is upper semicontinuous with nonempty, compact, convex values and locally bounded, the existence of Filippov solutions of (3) is guaranteed (cf. [10]). A *maximal solution* is a Filippov solution whose domain of existence is maximal, i.e., cannot be extended any further. A set M is *weakly invariant* (resp. *strongly invariant*) for (3) if for each $x_0 \in M$, M contains a maximal solution (resp. all maximal solutions) of (3).

2.2 Stability via nonsmooth Lyapunov functions

Let $f : \mathbb{R}^d \to \mathbb{R}$ be locally Lipschitz and regular (see [4] for detailed definitions). From Rademacher's Theorem [4], we know that locally Lipschitz functions are differentiable a.e. Let $\Omega_f \subset \mathbb{R}^d$ denote the set of points where f fails to be differentiable. The *generalized gradient of f at $x \in \mathbb{R}^d$* (cf. [4]) is defined by

$$\partial f(x) = \operatorname{co}\{\lim_{i \to +\infty} df(x_i) \mid x_i \to x\,,\ x_i \notin S \cup \Omega_f\},$$

where S can be any set of zero measure. Note that if f is continuously differentiable, then $\partial f(x) = \{df(x)\}$.

Given a locally Lipschitz function f, the *set-valued Lie derivative of f with respect to X* at x (cf. [1, 8]) is

$$\widetilde{\mathcal{L}}_X f(x) = \{a \in \mathbb{R} \mid \exists v \in K[X](x) \text{ with } \zeta \cdot v = a\,,\ \forall \zeta \in \partial f(x)\}.$$

For $x \in \mathbb{R}^d$, $\widetilde{\mathcal{L}}_X f(x)$ is a closed and bounded interval in $\mathbb{R}$, possibly empty. If f is continuously differentiable at x, then $\widetilde{\mathcal{L}}_X f(x) = \{df \cdot v \mid v \in K[X](x)\}$. If, in addition, X is continuous at x, then $\widetilde{\mathcal{L}}_X f(x) = \{\mathcal{L}_X f(x)\}$, the usual Lie derivative of f in the direction of X at x. The next result, from [1], shows that this Lie derivative measures the evolution of a function along the Filippov solutions.

Theorem 1 *Let $x : [t_0, t_1] \to \mathbb{R}^d$ be a Filippov solution of (3). Let f be a locally Lipschitz and regular function. Then $t \mapsto f(x(t))$ is absolutely continuous, $\frac{d}{dt}\big(f(x(t))\big)$ exists a.e. and $\frac{d}{dt}\left(f(x(t))\right) \in \widetilde{\mathcal{L}}_X f(x(t))$ a.e.*

Sometimes, we can also look at second-order information for the evolution of a function along the Filippov solutions. This is what we prove in the next result.

Proposition 2 *Let $x : [t_0, t_1] \to \mathbb{R}^d$ be a Filippov solution of (3). Let f be a locally Lipschitz and regular function. Assume that $\widetilde{\mathcal{L}}_X f : \mathbb{R}^d \to 2^{\mathbb{R}}$ is single-valued, i.e., it takes the form $\widetilde{\mathcal{L}}_X f : \mathbb{R}^d \to \mathbb{R}$, and assume it is Lipschitz and regular. Then $\frac{d^2}{dt^2}\left(f(x(t))\right)$ exists a.e. and $\frac{d^2}{dt^2}\left(f(x(t))\right) \in \widetilde{\mathcal{L}}_X(\widetilde{\mathcal{L}}_X f)(x(t))$ a.e.*

PROOF. Applying Theorem 1 to f and $\widetilde{\mathcal{L}}_X f$, resp., we deduce that (i) $t \mapsto f(x(t))$ is absolutely continuous, and $\frac{d}{dt}\big(f(x(t))\big) = \widetilde{\mathcal{L}}_X f(x(t))$ a.e., and, (ii) $t \mapsto \widetilde{\mathcal{L}}_X f(x(t))$ is absolutely continuous, and $\frac{d}{dt}\big(\widetilde{\mathcal{L}}_X f(x(t))\big) = \widetilde{\mathcal{L}}_X(\widetilde{\mathcal{L}}_X f)(x(t))$ a.e. Since $t \mapsto \widetilde{\mathcal{L}}_X f(x(t))$ is continuous, the expression

$$\begin{aligned} f(x(t)) &= f(x(t_0)) + \int_{t_0}^{t} \frac{d}{dt}\left(f(x(s))\right) ds \\ &= f(x(t_0)) + \int_{t_0}^{t} \widetilde{\mathcal{L}}_X f(x(s)) ds, \end{aligned}$$

and the second fundamental theorem of calculus implies that $t \mapsto f(x(t))$ is continuously differentiable, and therefore $\frac{d}{dt}\left(f(x(t))\right) = \widetilde{\mathcal{L}}_X f(x(t))$ for all t. Now, using (ii), we conclude that $t \mapsto \frac{d}{dt}\left(f(x(t))\right)$ is differentiable a.e. and $\frac{d^2}{dt^2}\left(f(x(t))\right) \in \widetilde{\mathcal{L}}_X(\widetilde{\mathcal{L}}_X f)(x(t))$ a.e. •

The following result is a generalization of LaSalle principle for discontinuous differential equations (3) with nonsmooth Lyapunov functions. The formulation is taken from [1], and slightly generalizes [19].

Theorem 3 (LaSalle Invariance Principle): *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally Lipschitz and regular function. Let $x_0 \in S \subset \mathbb{R}^d$, with S compact and strongly invariant for (3). Assume that either $\max \widetilde{\mathcal{L}}_X f(x) \leq 0$ or $\widetilde{\mathcal{L}}_X f(x) = \emptyset$ for all $x \in S$. Let $Z_{X,f} = \{x \in \mathbb{R}^d \mid 0 \in \widetilde{\mathcal{L}}_X f(x)\}$. Then, any solution $x : [t_0, +\infty) \to \mathbb{R}^d$ of (3) starting from x_0 converges to the largest weakly invariant set M contained in $\overline{Z}_{X,f} \cap S$. Moreover, if the set M is a finite collection of points, then the limit of all solutions starting at x_0 exists and equals one of them.*

The following result is taken from [8].

Proposition 4 (Finite-time convergence with first-order information): *Under the same assumptions of Theorem 3, further assume that there exists a neighborhood U of $Z_{X,f} \cap S$ in S such that $\max \widetilde{\mathcal{L}}_X f \leq -\epsilon < 0$ a.e. on $U \setminus (Z_{X,f} \cap S)$. Then, any solution $x : [t_0, +\infty) \to \mathbb{R}^d$ of (3) starting at $x_0 \in S$ reaches $Z_{X,f} \cap S$ in finite time. Moreover, if $U = S$, then the convergence time is upper bounded by $(f(x_0) - \min_{x \in S} f(x))/\epsilon$.*

Often times, first-order information is inconclusive to assess finite-time convergence. The next result uses second-order information to arrive at a satisfactory answer.

Theorem 5 (Finite-time convergence with second-order information): *Under the same assumptions of Theorem 3, further assume that (i) $x \in \mathbb{R}^d \mapsto \widetilde{\mathcal{L}}_X f(x)$ is single-valued, Lipschitz and regular; and (ii) there exists a neighborhood U of $Z_{X,f} \cap S$ in S such that $\min \widetilde{\mathcal{L}}_X(\widetilde{\mathcal{L}}_X f) \geq \epsilon > 0$ a.e. on $U \setminus (Z_{X,f} \cap S)$. Then, any solution $x : [t_0, +\infty) \to \mathbb{R}^d$ of (3) starting at $x_0 \in S$ reaches $Z_{X,f} \cap S$ in finite time. Moreover, if $U = S$, then the convergence time is upper bounded by $-\widetilde{\mathcal{L}}_X f(x_0)/\epsilon$.*

PROOF. Since $x \in \mathbb{R}^d \mapsto \widetilde{\mathcal{L}}_X f(x)$ is single-valued and continuous, then $Z_{X,f} = \{x \in \mathbb{R}^d \mid \widetilde{\mathcal{L}}_X f(x) = 0\}$ is closed. Let $x : [t_0, +\infty) \to \mathbb{R}^d$ be a solution of (3) starting from $x_0 \in S \setminus Z_{X,f}$. We reason by contradiction. Assume it does not exist T such that $x(T) \in Z_{X,f}$. By Theorem 3, $x(t) \to M \subset Z_{X,f} \cap S$ when $t \to +\infty$, and therefore there exists $t_* \geq t_0$ with $x(t) \in U$ for all $t \geq t_*$. Using Proposition 2 (cf. assumption (i)) combined with assumption (ii), we write for $g(t) = \frac{d}{dt}\big(f(x(t))\big)$,

$$g(t) = g(t_*) + \int_{t_*}^{t} \frac{d}{ds} g(s) ds \geq g(t_*) + \epsilon(t - t_*), \; t > t_*.$$

Since $x(t_*) \notin Z_{X,f}$ by hypothesis, then $g(t_*) < 0$. Noting that $t \mapsto g(t)$ is continuous, we deduce that there exists $T \leq t_* - \frac{g(t_*)}{\epsilon}$ such that $g(T) = 0$, i.e., $x(T) \in Z_{X,f}$, which is a contradiction. The upper bound on the convergence time can be deduced using similar arguments. •

3 Finite-time convergent gradient flows

Here, we formally introduce the normalized and signed gradient flows of a differentiable function, and characterize their convergence properties. We build on Section 2 to identify conditions for finite-time convergence. Let

$$\dot{x} = -\frac{\operatorname{grad}(f)(x)}{\|\operatorname{grad}(f)(x)\|_2}, \tag{6a}$$

$$\dot{x} = -\operatorname{sgn}(\operatorname{grad}(f)(x)). \tag{6b}$$

Both equations have discontinuous right-hand sides. Hence, we understand their solutions in the Filippov sense. Note that the trajectories of (6a) and of $\dot{x} = -\operatorname{grad}(f)(x)$ describe the same paths.

Lemma 6 *The Filippov set-valued maps associated with the discontinuous vector fields* (6a) *and* (6b) *are*

$$K\Big[\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}\Big](x) = \operatorname{co}\Big\{\lim_{i\to+\infty}\frac{\operatorname{grad}(f)(x_i)}{\|\operatorname{grad}(f)(x_i)\|_2} \mid x_i \to x,\ \operatorname{grad}(f)(x_i)\neq 0\Big\},$$

$$K\big[\operatorname{sgn}(\operatorname{grad}(f))\big](x) = \big\{v\in\mathbb{R}^d \mid v_i = \operatorname{sgn}(\operatorname{grad}_i(f)(x)) \text{ if } \operatorname{grad}_i(f)(x)\neq 0 \text{ and } v_i\in[-1,1] \text{ if } \operatorname{grad}_i(f)(x)=0, \text{ for } i\in\{1,\dots,d\}\big\}.$$

Note $K\big[\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}\big](x) = \frac{\operatorname{grad}(f)(x)}{\|\operatorname{grad}(f)(x)\|_2}$ *if* $\operatorname{grad}(f)(x)\neq 0$.

The proof of this result follows from the definition (4) of the operator K and the particular forms of (6a) and (6b).

For a differentiable function f, let $\operatorname{Critical}(f) = \big\{x\in\mathbb{R}^d \mid \operatorname{grad}(f)(x)=0\big\}$ denote the set of its critical points, and let $\operatorname{Hess}(f)(x)$ denote its Hessian matrix at $x\in\mathbb{R}^d$. The next result establishes the general asymptotic properties of the flows in (6).

Proposition 7 *Let $f:\mathbb{R}^d\to\mathbb{R}$ be a differentiable function. Let $x_0\in S\subset\mathbb{R}^d$, with S compact and strongly invariant for* (6a) *(resp., for* (6b)*). Then each solution of equation* (6a) *(resp. equation* (6b)*) starting from x_0 asymptotically converges to* Critical(f).

PROOF. For equation (6a), if $\operatorname{grad}(f)(x)\neq 0$, then

$$\widetilde{\mathcal{L}}_{\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}} f(x) = \Big\{\frac{\operatorname{grad}(f)(x)}{\|\operatorname{grad}(f)(x)\|_2}\cdot\operatorname{grad}(f)(x)\Big\} = \big\{\|\operatorname{grad}(f)(x)\|_2\big\}.$$

If, instead, $\operatorname{grad}(f)(x)=0$, then $\widetilde{\mathcal{L}}_{\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}} f(x) = \{0\}$. Therefore, we deduce

$$\widetilde{\mathcal{L}}_{-\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}} f(x) = -\|\operatorname{grad}(f)(x)\|_2, \quad \text{for all } x\in\mathbb{R}^d.$$

Consequently, $Z_{-\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2},f} = \operatorname{Critical}(f)$ is closed, and LaSalle Invariance Principle (cf. Theorem 3) implies that each solution of (6a) starting from x_0 asymptotically converges to the largest weakly invariant set M contained in $\operatorname{Critical}(f)\cap S$, which is $\operatorname{Critical}(f)\cap S$ itself.

For equation (6b), we have $a\in\widetilde{\mathcal{L}}_{\operatorname{sgn}(\operatorname{grad}(f))}f(x)$ if and only if there exists $v\in K\big[\operatorname{sgn}(\operatorname{grad}(f))\big](x)$ such that $a = v\cdot\operatorname{grad}(f)(x)$. From Lemma 6, we deduce that $a = \operatorname{sgn}(\operatorname{grad}_1(f)(x))\cdot\operatorname{grad}_1(f)(x)+\dots+\operatorname{sgn}(\operatorname{grad}_n(f)(x))\cdot\operatorname{grad}_n(f)(x) = \|\operatorname{grad}(f)(x)\|_1$. Therefore, we deduce

$$\widetilde{\mathcal{L}}_{-\operatorname{sgn}(\operatorname{grad}(f))}f(x) = \{-\|\operatorname{grad}(f)(x)\|_1\}.$$

Consequently, $Z_{-\operatorname{sgn}(\operatorname{grad}(f)),f} = \operatorname{Critical}(f)$ is closed, and LaSalle Invariance Principle implies that each solution of (6b) starting from x_0 asymptotically converges to the largest weakly invariant set M contained in $\operatorname{Critical}(f)\cap S$, which is $\operatorname{Critical}(f)\cap S$ itself. •

Let us now discuss the finite-time convergence properties of the vector fields (6). Note that Proposition 4 cannot be applied. Indeed,

$$\max\widetilde{\mathcal{L}}_{-\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}} f(x) = -\|\operatorname{grad}(f)(x)\|_2,$$
$$\max\widetilde{\mathcal{L}}_{-\operatorname{sgn}(\operatorname{grad}(f))} f(x) = -\|\operatorname{grad}(f)(x)\|_1,$$

and both $\inf_{x\in U\setminus\operatorname{Critical}(f)\cap S}\|\operatorname{grad}(f)(x)\|_2 = 0$ and $\inf_{x\in U\setminus\operatorname{Critical}(f)\cap S}\|\operatorname{grad}(f)(x)\|_1 = 0$, for any neighborhood U of $\operatorname{Critical}(f)\cap S$ in S. Hence, the hypotheses of Proposition 4 are not verified by either (6a) or (6b).

Under additional conditions, one can establish stronger convergence properties of (6). We show this next.

Theorem 8 *Let $f:\mathbb{R}^d\to\mathbb{R}$ be a second-order differentiable function. Let $x_0\in S\subset\mathbb{R}^d$, with S compact and strongly invariant for* (6a) *(resp., for* (6b)*). Assume there exists a neighborhood V of* $\operatorname{Critical}(f)\cap S$ *in S where either one of the following conditions hold:*

(i) for all $x\in V$, $\operatorname{Hess}(f)(x)$ *is positive definite; or*
(ii) for all $x\in V\setminus(\operatorname{Critical}(f)\cap S)$, $\operatorname{Hess}(f)(x)$ *is positive semidefinite, the multiplicity of the eigenvalue* 0 *is constant, and* $\operatorname{grad}(f)(x)$ *is orthogonal to the eigenspace of* $\operatorname{Hess}(f)(x)$ *corresponding to* 0.

Then each solution of (6a) *(resp.* (6b)*) starting from x_0 converges in finite time to a critical point of f. Furthermore, if $V=S$, then the convergence time of the solutions of* (6a) *(resp.* (6b)*) starting from x_0 is upper bounded by*

$$\frac{1}{\lambda_0}\|\operatorname{grad}(f)(x_0)\|_2 \quad \Big(\textit{resp. } \frac{1}{\lambda_0}\|\operatorname{grad}(f)(x_0)\|_1\Big),$$

where $\lambda_0 = \min_{x\in S}\lambda_2(\operatorname{Hess}(f)(x))$.

PROOF. Our strategy is to show that the hypotheses of Theorem 5 are verified by both vector fields. From Proposition 7, we know that each solution of (6a) (resp. (6b)) starting from x_0 converges to $\operatorname{Critical}(f)$. Let us take an open set $U\subset S$ such that $\operatorname{Critical}(f)\cap S\subset U\subset\overline{U}\subset V$. Since S is compact, $\overline{U}$ is also compact. By continuity, under either assumption (i) or assumption (ii), the function $\lambda_2(\operatorname{Hess}(f)):\overline{U}\to\mathbb{R}$, $x\mapsto\lambda_2(\operatorname{Hess}(f)(x))$, reaches its minimum on $\overline{U}$, i.e, there exists $\lambda_0>0$ such that $\lambda_2(\operatorname{Hess}(f)(x))\geq\lambda_0$ for all $x\in\overline{U}$. Moreover, from (1), we have for all $u\in\mathbb{R}^d$,

$$u'\operatorname{Hess}(f)(x)\,u \geq \geq \lambda_2(\operatorname{Hess}(f)(x))\,\|u-\pi_{H_0(\operatorname{Hess}(f)(x))}(u)\|_2^2. \quad (7)$$

For (6a), recall from the proof of Proposition 7, that the function $x\in\mathbb{R}^d\mapsto\widetilde{\mathcal{L}}_{\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}} f(x) = \|\operatorname{grad}(f)(x)\|_2$ is single-valued, locally Lipschitz and regular, and hypothesis (i) in Theorem 5 is satisfied. Additionally,

$Z_{-\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2},f} = \operatorname{Critical}(f)$. Let us take $x \notin \operatorname{Critical}(f)$, and let us compute $\widetilde{\mathcal{L}}_{\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}}(\widetilde{\mathcal{L}}_{\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}} f)(x)$. Noting

$$\partial(\|\operatorname{grad}(f)\|_2)(x) = \left\{ \operatorname{Hess}(f)(x) \frac{\operatorname{grad}(f)(x)}{\|\operatorname{grad}(f)(x)\|_2} \right\},$$

we deduce

$$\widetilde{\mathcal{L}}_{\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}}(\widetilde{\mathcal{L}}_{\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}} f)(x) = \frac{\operatorname{grad}(f)(x)'}{\|\operatorname{grad}(f)(x)\|_2} \operatorname{Hess}(f)(x) \frac{\operatorname{grad}(f)(x)}{\|\operatorname{grad}(f)(x)\|_2}. \quad (8)$$

Let $x \in \overline{U} \setminus (\operatorname{Critical}(f) \cap S)$. Under either assumption (i) or (ii) in the theorem, $\pi_{H_0(\operatorname{Hess} f(x))}\big(\frac{\operatorname{grad}(f)(x)}{\|\operatorname{grad}(f)(x)\|_2}\big) = 0$. Then, using (7) in equation (8), we conclude

$$\widetilde{\mathcal{L}}_{\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}}(\widetilde{\mathcal{L}}_{\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}} f)(x) \geq \lambda_2(\operatorname{Hess}(f)(x)) \cdot \cdot \left\| \frac{\operatorname{grad}(f)(x)}{\|\operatorname{grad}(f)(x)\|_2} \right\|_2^2 = \lambda_2(\operatorname{Hess}(f)(x)) \geq \lambda_0 > 0,$$

for $x \in \overline{U} \setminus (\operatorname{Critical}(f) \cap S)$. Hence, hypothesis (ii) in Theorem 5 is also verified, and we deduce that the set $\operatorname{Critical}(f)$ is reached in finite time, which in particular implies that the limit of any solution of equation (6a) starting from $x_0 \in S$ exists and is reached in finite time.

For (6b), recall from Proposition 7, that the function $x \in \mathbb{R}^d \mapsto \widetilde{\mathcal{L}}_{\operatorname{sgn}(\operatorname{grad}(f))} f(x) = \|\operatorname{grad}(f)(x)\|_1$ is single-valued, locally Lipschitz and regular, and hypothesis (i) in Theorem 5 is satisfied. Additionally, $Z_{-\operatorname{sgn}(\operatorname{grad}(f)),f} = \operatorname{Critical}(f)$. Let us take $x \notin \operatorname{Critical}(f)$, and compute $\widetilde{\mathcal{L}}_{\operatorname{sgn}(\operatorname{grad}(f))}(\widetilde{\mathcal{L}}_{\operatorname{sgn}(\operatorname{grad}(f))} f)(x)$. By definition, $a \in \widetilde{\mathcal{L}}_{\operatorname{sgn}(\operatorname{grad}(f))}(\widetilde{\mathcal{L}}_{\operatorname{sgn}(\operatorname{grad}(f))} f)(x)$ if and only if there exists $v \in K\big[\operatorname{sgn}(\operatorname{grad}(f))\big](x)$ such that $a = v \cdot \zeta$, for any $\zeta \in \partial(\|\operatorname{grad}(f)\|_1)(x)$. Note that

$$\begin{aligned}
&\partial(\|\operatorname{grad}(f)\|_1)(x) = \\
&\big\{\zeta \in \mathbb{R}^d \mid \zeta = \operatorname{Hess}(f)(x)\,\eta, \text{ for some } \eta \in \mathbb{R}^d \text{ with} \\
&\quad \eta_i = \operatorname{sgn}(\operatorname{grad}_i(f)(x)) \text{ if } \operatorname{grad}_i(f)(x) \neq 0 \text{ and} \\
&\quad \eta_i \in [-1,1] \text{ if } \operatorname{grad}_i(f)(x) = 0, \text{ for } i \in \{1,\dots,d\}\big\}.
\end{aligned}$$

In particular, $\operatorname{Hess}(f)(x)\, v \in \partial(\|\operatorname{grad}(f)\|_1)(x)$. Then $a = v' \operatorname{Hess}(f)(x)\, v$. Let us now decompose v as $v = \pi_{H_0(x)}(v) + (v - \pi_{H_0(x)}(v))$, where $\pi_{H_0(x)}(v) \in H_0(x)$ and $v - \pi_{H_0(x)}(v) \in H_0(x)^\perp$. Because $v \in K\big[\operatorname{sgn}(\operatorname{grad}(f))\big](x)$, we deduce $v \cdot \operatorname{grad}(f)(x) = \|\operatorname{grad}(f)(x)\|_1$. Let $x \in \overline{U} \setminus (\operatorname{Critical}(f) \cap S)$. Under either assumption (i) or (ii),

$$\begin{aligned}
\|\operatorname{grad}(f)(x)\|_1 &= v \cdot \operatorname{grad}(f)(x) \\
&= (v - \pi_{H_0(x)}(v)) \cdot \operatorname{grad}(f)(x) \\
&\leq \|v - \pi_{H_0(x)}(v)\|_2 \|\operatorname{grad}(f)(x)\|_2.
\end{aligned}$$

Using $\|u\|_1 \geq \|u\|_2$ for any $u \in \mathbb{R}^d$, we deduce from this equation that $\|v - \pi_{H_0(x)}(v)\|_2 \geq 1$. Therefore, using (7)

$$\begin{aligned}
a = v' \operatorname{Hess}(f)(x)\, v &\geq \lambda_2(\operatorname{Hess}(f)(x))\, \|v - \pi_{H_0(x)}(v)\|_2^2 \\
&\geq \lambda_2(\operatorname{Hess}(f)(x)) \geq \lambda_0 > 0,
\end{aligned}$$

for $x \in \overline{U} \setminus (\operatorname{Critical}(f) \cap S)$. Consequently, we get $\min \widetilde{\mathcal{L}}_{\operatorname{sgn}(\operatorname{grad}(f))}(\widetilde{\mathcal{L}}_{\operatorname{sgn}(\operatorname{grad}(f))} f) \geq \lambda_0 > 0$ on $\overline{U} \setminus (\operatorname{Critical}(f) \cap S)$. Hence, hypothesis (ii) in Theorem 5 is also verified, and we deduce that the set $\operatorname{Critical}(f)$ is reached in finite time, which in particular implies that the limit of any solution of equation (6b) starting from $x_0 \in S$ exists and is reached in finite time. The upper bounds on the convergence time of the solutions of both flows also follow from Theorem 5. •

Corollary 9 *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a second-order differentiable function. Let $x_0 \in S \subset \mathbb{R}^d$, with S compact and strongly invariant for* (6a) *(resp., for* (6b)*). Assume that for each $x \in \operatorname{Critical}(f) \cap S$, $\operatorname{Hess}(f)(x)$ is positive definite. Then each solution of* (6a) *(resp.* (6b)*) starting from x_0 converges in finite time to a minimum of f.*

4 Applications to network consensus

The results of the preceding sections on nonsmooth gradient dynamical systems can be applied to any multi-agent coordination algorithm whose design involves the gradient of meaningful aggregate objective functions. As an illustration, we discuss here the application to network consensus problems, and defer the treatment of other coordination tasks to future work.

Consider a network of n agents. The state of the ith agent, denoted $p_i \in \mathbb{R}$, evolves according to a first-order dynamics $\dot{p}_i(t) = u_i$. Let $G = (\{1,\dots,n\}, E)$ be an undirected graph with n vertices, describing the topology of the network. The graph Laplacian matrix L_G associated with G (see, for instance, [16]) is defined as $L_G = \Delta_G - A_G$, where Δ_G is the degree matrix and A_G is the adjacency matrix of the graph. When the graph G is clear from the context, we will simply denote it by L. The Laplacian matrix is symmetric, positive semidefinite and has 0 as an eigenvalue with eigenvector $\mathbf{1}$. More importantly, G is connected if and only if $\operatorname{rank}(L) = n-1$, i.e., if the eigenvalue 0 has multiplicity one. This is the reason why $\lambda_2(L) = \min\{\lambda \mid \lambda > 0 \text{ and } \lambda \text{ eigenvalue of } L\}$ is termed the *algebraic connectivity* of G.

Two agents p_i and p_j *agree* if and only if $p_i = p_j$. The *Laplacian potential* $\Phi_G : \mathbb{R}^n \to \overline{\mathbb{R}}_+$ associated with G (see [16]) quantifies the group disagreement,

$$\Phi_G(p_1,\dots,p_n) = \frac{1}{2} P' L P = \frac{1}{2} \sum_{(i,j)\in E} (p_j - p_i)^2,$$

with $P = (p_1,\dots,p_n)' \in \mathbb{R}^n$. Clearly, $\Phi_G(p_1,\dots,p_n) = 0$ if and only if all neighboring nodes in the graph G agree. If G is connected, then all nodes agree and a consensus is reached. Therefore, we want the network to reach the critical points of Φ_G. Assume G is connected. The Laplacian potential is smooth, and its gradient is $\mathrm{grad}(\Phi_G)(P) = LP$. The gradient algorithm

$$\dot{p}_i(t) = -\frac{\partial \Phi_G}{\partial p_i} = \sum_{j \in \mathcal{N}_{G,i}} (p_j(t) - p_i(t)), \tag{9}$$

for $i \in \{1, \dots, n\}$, is distributed over G, i.e., each agent can implement it with the information provided by its neighbors in the graph G (see [9] for a more thorough exposition of *spatially distributed* algorithms). The algorithm (9) asymptotically converges to the critical points of Φ_G, i.e., asymptotically achieves consensus. Actually, since the system is linear, the convergence is exponential with rate lower bounded by $\lambda_2(L)$. Additionally, the fact that $\mathbf{1} \cdot (LP) = 0$ implies that $\sum_{i=1}^n p_i$ is constant along the solutions. Therefore, each solution of (9) is convergent to a point of the form $(p_*, \dots, p_*)$, with $p_* = \frac{1}{n} \sum_{i=1}^n p_i(0)$ (this is called *average-consensus*).

Following (6), consider the discontinuous algorithms

$$\dot{p}_i(t) = \frac{\sum_{j \in \mathcal{N}_{G,i}} (p_j(t) - p_i(t))}{\|LP(t)\|_2}, \tag{10a}$$

$$\dot{p}_i(t) = \mathrm{sgn}\big(\sum_{j \in \mathcal{N}_{G,i}} (p_j(t) - p_i(t))\big), \tag{10b}$$

for $i \in \{1, \dots, n\}$. Note that the algorithm (10b) is distributed over G, whereas the algorithm (10a) is not. Before analyzing their convergence properties, we identify a conserved quantity for each of these flows.

Proposition 10 *Define $g_1 : \mathbb{R}^n \to \mathbb{R}$, $g_2 : \mathbb{R}^n \to \mathbb{R}$ by*

$$g_1(p_1, \dots, p_n) = \sum_{i=1}^n p_i,$$
$$g_2(p_1, \dots, p_n) = \max_{i \in \{1,\dots,n\}} \{p_i\} + \min_{i \in \{1,\dots,n\}} \{p_i\}.$$

Then g_1 is constant along the solutions of (10a) *and g_2 is constant along the solutions of* (10b)*.*

PROOF. The function g_1 is differentiable, with $\mathrm{grad}(g_1)(P) = \mathbf{1}$. For any $P = (p_1, \dots, p_n) \in \mathbb{R}^n$, $\widetilde{\mathcal{L}}_{-\frac{\mathrm{grad}(\Phi_G)}{\|\mathrm{grad}(\Phi_G)\|_2}} g_1(P) = \{0\}$. Therefore, from Theorem 1, we conclude that g_1 is constant along the solutions of (10a). On the other hand, from [4, Proposition 2.3.12], one deduces that g_2 is locally Lipschitz and regular, with

$$\partial g_2(P) = \mathrm{co}\big\{e_j \in \mathbb{R}^n \mid j \text{ such that } p_j = \min_{i \in \{1,\dots,n\}} \{p_i\}\big\} + \mathrm{co}\big\{e_k \in \mathbb{R}^n \mid k \text{ such that } p_k = \max_{i \in \{1,\dots,n\}} \{p_i\}\big\}.$$

Let $a \in \widetilde{\mathcal{L}}_{-\mathrm{sgn}(\mathrm{grad}(\Phi_G))} g_2(P)$. By definition, there exists $v \in K\big[-\mathrm{sgn}(\mathrm{grad}(\Phi_G))\big](P)$ with

$$a = v \cdot \zeta, \quad \text{for all } \zeta \in \partial g_2(P). \tag{11}$$

If $P \in \mathrm{diag}(\mathbb{R}^n)$, then $\partial g_2(P) = \mathbb{R}^d$, and, for (11) to hold, necessarily $v = (0, \dots, 0)$. Therefore, $a = 0$. If $P \notin \mathrm{diag}(\mathbb{R}^n)$, there exist $j, k \in \{1, \dots, n\}$ with $p_j = \min_{i \in \{1,\dots,n\}} \{p_i\}$, $p_k = \max_{i \in \{1,\dots,n\}} \{p_i\}$ such that

$$\sum_{i \in \mathcal{N}_{G,j}} (p_i - p_j) > 0, \quad \sum_{i \in \mathcal{N}_{G,k}} (p_i - p_k) < 0,$$

and hence, from Lemma 6, $v_j = 1$ and $v_k = -1$. Therefore, we deduce $a = v \cdot (e_j + e_k) = 1 - 1 = 0$. Note that $\widetilde{\mathcal{L}}_{-\mathrm{sgn}(\mathrm{grad}(\Phi_G))} g_2(P) \neq \emptyset$ because $\mathrm{sgn}(\mathrm{grad}(\Phi_G)) \cdot \zeta = 0$ for all $\zeta \in \partial g_2(P)$, and hence $0 \in \widetilde{\mathcal{L}}_{-\mathrm{sgn}(\mathrm{grad}(\Phi_G))} g_2(P)$. Finally, we conclude $\widetilde{\mathcal{L}}_{-\mathrm{sgn}(\mathrm{grad}(\Phi_G))} g_2(P) = \{0\}$, and therefore g_2 is constant along the solutions of (10b). •

The following theorem completely characterizes the asymptotic convergence properties of the flows in (10).

Theorem 11 *Let $G = (\{1, \dots, n\}, E)$ be a connected undirected graph. Then, the flows in* (10) *achieve consensus in finite time. More specifically, for $P_0 = ((p_1)_0, \dots, (p_n)_0) \in \mathbb{R}^n$,*

(i) the solutions of (10a) *starting from P_0 converge in finite time to $(p_*, \dots, p_*)$, with $p_* = \frac{1}{n} \sum_{i=1}^n (p_i)_0$ (average-consensus). The convergence time is upper bounded by $\|LP_0\|_2 / \lambda_2(L)$;*

(ii) the solutions of (10b) *starting from P_0 converge in finite time to $(p_*, \dots, p_*)$, with $p_* = \frac{1}{2}\big(\max_{i \in \{1,\dots,n\}} \{(p_i)_0\} + \min_{i \in \{1,\dots,n\}} \{(p_i)_0\}\big)$ (average-max-min-consensus). The convergence time is upper bounded by $\|LP_0\|_1 / \lambda_2(L)$.*

PROOF. Our strategy is to verify the assumptions in Theorem 8. Let $\Phi_G^{-1}(\leq \Phi_G(P_0)) = \{(p_1, \dots, p_n) \in \mathbb{R}^d \mid \Phi_G(p_1, \dots, p_n) \leq \Phi_G(P_0)\}$. Clearly, this set is strongly invariant for both flows. Since L is positive semidefinite, $\Phi_G(p_1, \dots, p_n) \geq \lambda_2(L) \, \|P - \pi_{H_0(A)}(P)\|_2^2$. Then, $\|P - \pi_{H_0(A)}(P)\|_2^2 \leq \Phi_G(P_0)/\lambda_2(L)$ for $P \in \Phi_G^{-1}(\leq \Phi_G(P_0))$. Consider also the closed set

$$W(P_0) = \big\{P \in \mathbb{R}^n \mid \min_{i\{1,\dots,n\}} \{(p_i)_0\} \leq \frac{1}{n} P \cdot \mathbf{1} \leq \max_{i \in \{1,\dots,n\}} \{(p_i)_0\}\big\}.$$

One can see that $W(P_0)$ is strongly invariant for (10a) and for (10b). Now, define the set $S = W(P_0) \cap \Phi_G^{-1}(\leq \Phi_G(P_0))$. From the preceding discussion, we deduce that S is strongly invariant for (10a) and (10b). Clearly, S is closed. Furthermore, using $P = \pi_{H_0(L)}(P) + P - \pi_{H_0(L)}(P)$, and noting $\pi_{H_0(L)}(P) = \frac{P \cdot \mathbf{1}}{n} \mathbf{1}$, we deduce

$$\|P\|_2 = \|\pi_{H_0(L)}(P)\|_2 + \|P - \pi_{H_0(L)}(P)\|_2$$
$$\leq \sqrt{n} \max\big\{|\min_{i\{1,\dots,n\}} \{(p_i)_0\}|, |\max_{i\{1,\dots,n\}} \{(p_i)_0\}|\big\} + \frac{\Phi_G(P_0)}{\lambda_2(L)}$$

for $P \in S$. Therefore, S is bounded, and hence compact. Now, $\mathrm{Hess}(\Phi_G)(P) = L$ is positive semidefinite at any $P \in \mathbb{R}^n$, with the eigenvalue 0 having a multiplicity 1 (not depending on P). Additionally, for $P \notin \mathrm{Critical}(\Phi_G)$, $\mathrm{grad}(\Phi_G)(P) = LP \neq 0$ is orthogonal to $\mathrm{span}\{\mathbf{1}\}$, the eigenspace of L corresponding to the eigenvalue 0. Finally, Theorem 8(ii) with $V = S$ (together with Proposition 10) yields the result. •

Fig. 1 illustrates the evolution of the differential equations (9), (10a) and (10b). As stated in Theorem 11, the agents' states evolving under (10a) achieve consensus in finite time at $\frac{1}{n}\sum_{i=1}^{n}(p_i)_0$, and the agents' states evolving under (10b) achieve consensus in finite time at $\frac{1}{2}\big(\max_{i\in\{1,\dots,n\}}\{(p_i)_0\} + \min_{i\in\{1,\dots,n\}}\{(p_i)_0\}\big)$.

Networks with switching network topologies. For networks with switching connected topologies, one can derive a similar result to Theorem 11. Consider, following [16], the next setup. Let Γ_n be the finite set of connected undirected graphs with vertices $\{1,\dots,n\}$,

$$\Gamma_n = \{G = (\{1,\dots,n\}, E) \mid G \text{ connected, undirected}\}.$$

Let $I_{\Gamma_n} \subset \mathbb{N}$ be an index set associated with the elements of Γ_n. A *switching signal* σ is a map $\sigma : \overline{\mathbb{R}}_+ \to I_{\Gamma_n}$. For each time $t \in \overline{\mathbb{R}}_+$, the switching signal σ establishes the network graph $G_{\sigma(t)} \in \Gamma_n$ employed by the network agents. Now, consider a network subject to the switching network topology defined by σ and executing one of the coordination algorithms introduced above. In other words, consider the switching system

$$\dot{p}_i(t) = -\frac{\partial \Phi_{G_{\sigma(t)}}}{\partial p_i} = \sum_{j\in\mathcal{N}_{G_{\sigma(t)},i}} (p_j(t) - p_i(t)), \qquad (12)$$

for $i \in \{1,\dots,n\}$, and the switching systems

$$\dot{p}_i(t) = \frac{\sum_{j\in\mathcal{N}_{G_{\sigma(t)},i}}(p_j(t) - p_i(t))}{\|L_{G_{\sigma(t)}}P(t)\|_2}, \qquad (13a)$$

$$\dot{p}_i(t) = \mathrm{sgn}\big(\sum_{j\in\mathcal{N}_{G_{\sigma(t)},i}} (p_j(t) - p_i(t))\big). \qquad (13b)$$

The switching system (12) asymptotically achieves average-consensus for an arbitrary switching signal σ. Let $G_* \in \Gamma_n$ be such that

$$\frac{\lambda_2(L_{G_*})}{\lambda_n(L_{G_*})} = \min_{G\in\Gamma_n}\Big\{\frac{\lambda_2(L_G)}{\lambda_n(L_G)}\Big\}.$$

For the systems in (13), we have the next result.

Corollary 12 *Let $\sigma : \overline{\mathbb{R}}_+ \to I_{\Gamma_n}$ be a switching signal. Then, the flow* (13a) *achieves average-consensus in finite time upper bounded by $\frac{\lambda_n(L_{G_*})}{\lambda_2(L_{G_*})}\big\|P(0) - \frac{1}{n}\sum_{i=1}^{n}(p_i)_0\,\mathbf{1}\big\|_2$, and the flow* (13b) *achieves average-max-min-consensus in finite time equal to $\frac{1}{2}\big(\max_{i\in\{1,\dots,n\}}\{(p_i)_0\} - \min_{i\in\{1,\dots,n\}}\{(p_i)_0\}\big)$.*

PROOF. For the flow (13a), consider the candidate Lyapunov function $V_1 : \mathbb{R}^n \to \mathbb{R}$

$$V_1(P) = \frac{1}{2}\big\|P - \frac{1}{n}\sum_{i=1}^{n} p_i\,\mathbf{1}\big\|_2^2.$$

The first-order evolution of this function along the network trajectories is determined by $\widetilde{\mathcal{L}}_{-\frac{\mathrm{grad}(\Phi_G)}{\|\mathrm{grad}(\Phi_G)\|_2}} V_1(P) = -\frac{P' L_G P}{\|L_G P\|_2}$, for each $G \in \Gamma_n$, which is single-valued, Lipschitz and regular. Additionally, for any $G \in \Gamma_n$,

$$\widetilde{\mathcal{L}}_{-\frac{\mathrm{grad}(\Phi_G)}{\|\mathrm{grad}(\Phi_G)\|_2}} V_1(P) \leq -\frac{\lambda_2(L_G)}{\lambda_n(L_G)}\big\|P - \frac{1}{n}\sum_{i=1}^{n} p_i\,\mathbf{1}\big\|_2 \leq 0.$$

The application of the LaSalle Invariance Principle ensures that the flow (13a) achieves average-consensus. From the preceding inequality and the definition of V_1,

$$\big\|P(t) - \frac{1}{n}\sum_{i=1}^{n} p_i\,\mathbf{1}\big\|_2 \leq \big\|P(0) - \frac{1}{n}\sum_{i=1}^{n} p_i\,\mathbf{1}\big\|_2 - \frac{\lambda_2(L_{G_*})}{\lambda_n(L_{G_*})}t,$$

which implies the result.

For the flow (13b), consider the candidate Lyapunov function $V_2 : \mathbb{R}^n \to \mathbb{R}$

$$V_2(P) = \big\|P - \frac{1}{2}\big(\max_{i\in\{1,\dots,n\}}\{p_i\} + \min_{i\in\{1,\dots,n\}}\{p_i\}\big)\,\mathbf{1}\big\|_\infty = \frac{1}{2}\big(\max_{i\in\{1,\dots,n\}}\{p_i\} - \min_{i\in\{1,\dots,n\}}\{p_i\}\big).$$

This function is locally Lipschitz and regular. Let $a \in \widetilde{\mathcal{L}}_{-\mathrm{sgn}(\mathrm{grad}(\Phi_G))}V_2(P)$. Then, there exists $v \in K\big[-\mathrm{sgn}(\mathrm{grad}(\Phi_G))\big](P)$ with $a = v \cdot \zeta$, for all $\zeta \in \partial V_2(P)$. Take P with $V_2(P) \neq 0$. Let $j, k \in \{1,\dots,n\}$ such that $p_j = \min_{i\in\{1,\dots,n\}}\{p_i\}$, $p_k = \max_{i\in\{1,\dots,n\}}\{p_i\}$. Then $\frac{1}{2}(e_k - e_j) \in \partial V_2(P)$, $v_j = 1$ and $v_k = -1$. Therefore $a = -1$. The result follows from Proposition 4. •

Remark 13 *Note that in the proof of Corollary 12, we have explicitly computed the convergence time of the flow* (10b) *to achieve average-max-min-consensus to be $\frac{1}{2}\big(\max_{i\in\{1,\dots,n\}}\{(p_i)_0\} - \min_{i\in\{1,\dots,n\}}\{(p_i)_0\}\big)$.* •

5 Conclusions

We have introduced the normalized and signed versions of the gradient descent flow of a differentiable function. We have characterized the asymptotic convergence properties of these nonsmooth gradient flows, and identified suitable conditions that guarantee that convergence to the critical points is achieved in finite time. In doing so, we have built on two novel nonsmooth analysis results on finite-time convergence and second-order information on the evolution of Lyapunov functions. These results are not restricted to gradient flows, and can indeed be used in other setups involving discontinuous vector fields. We

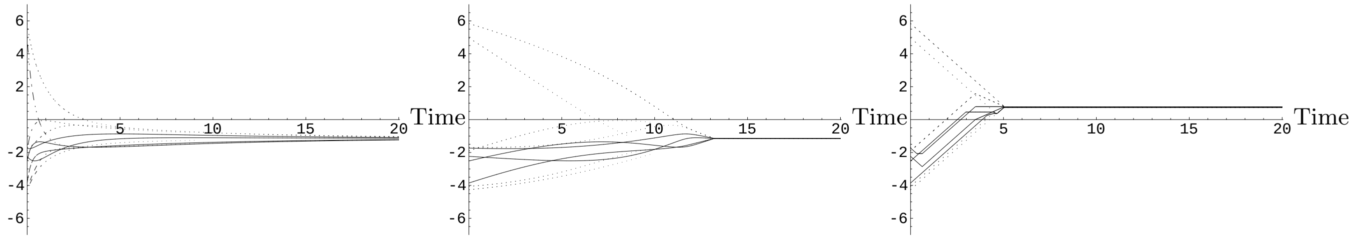

Fig. 1. From left to right, evolution of (9), (10a) and (10b) for 10 agents starting from a randomly generated initial configuration with $p_i \in [-7, 7]$, $i \in \{1, \ldots, 10\}$. The graph $G = (\{1, \ldots, 10\}, E)$ has edge set $E = \{(1, 4), (1, 10), (2, 10), (3, 6), (3, 9), (4, 8), (5, 6), (5, 9), (7, 10), (8, 9)\}$. The algebraic connectivity of G is $\lambda_2(L) = 0.12$.

have discussed the application of the results to network consensus problems.

Future work will be devoted to explore (i) the use of the upper bounds on the convergence time of the proposed nonsmooth flows in assessing the time complexity of coordination algorithms; (ii) the application of the results to consensus-based sensor fusion algorithms, and other coordination problems such as formation control, deployment and rendezvous; (iii) the robustness properties of the proposed consensus flows under asynchronism, delays and network topologies that are not connected at all times; and (iv) the identification of other nonsmooth distributed algorithms based on gradient information with similar convergence properties.

References

[1] A. Bacciotti and F. Ceragioli. Stability and stabilization of discontinuous systems and nonsmooth Lyapunov functions. *ESAIM. Control, Optimisation & Calculus of Variations*, 4: 361–376, 1999.

[2] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, Belmont, MA, 1997. ISBN 1886529019.

[3] S. P. Bhat and D. S. Bernstein. Finite-time stability of continuous autonomous systems. *SIAM Journal on Control and Optimization*, 38(3):751–766, 2000.

[4] F. H. Clarke. *Optimization and Nonsmooth Analysis*. Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley, 1983. ISBN 047187504X.

[5] F. H. Clarke. Lyapunov functions and feedback in nonlinear control. In M. S. de Queiroz, M. Malisoff, and P. Wolenski, editors, *Optimal Control, Stabilization and Nonsmooth Analysis*, volume 301 of *Lecture Notes in Control and Information Sciences*, pages 267–282. Springer Verlag, New York, 2004.

[6] F. H. Clarke, Y. Ledyaev, L. Rifford, and R. Stern. Feedback stabilization and Lyapunov functions. *SIAM Journal on Control and Optimization*, 39:25–48, 2000.

[7] J. M. Coron. On the stabilization in finite time of locally controllable systems by means of continuous time-varying feedback laws. *SIAM Journal on Control and Optimization*, 33:804–833, 1995.

[8] J. Cortés and F. Bullo. Coordination and geometric optimization via distributed dynamical systems. *SIAM Journal on Control and Optimization*, 44(5):1543–1574, 2005.

[9] J. Cortés, S. Martínez, and F. Bullo. Spatially-distributed coverage optimization and control with limited-range interactions. *ESAIM. Control, Optimisation & Calculus of Variations*, 11(4):691–719, 2005.

[10] A. F. Filippov. *Differential Equations with Discontinuous Righthand Sides*, volume 18 of *Mathematics and Its Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988.

[11] V. Gazi and K. M. Passino. Stability analysis of swarms. *IEEE Transactions on Automatic Control*, 48(4):692–697, 2003.

[12] U. Helmke and J.B. Moore. *Optimization and Dynamical Systems*. Springer Verlag, New York, 1994. ISBN 0387198571.

[13] W. M. Hirsch and S. Smale. *Differential Equations, Dynamical Systems and Linear Algebra*. Academic Press, New York, 1974.

[14] P. Ögren, E. Fiorelli, and N. E. Leonard. Cooperative control of mobile sensor networks: adaptive gradient climbing in a distributed environment. *IEEE Transactions on Automatic Control*, 49(8):1292–1302, 2004.

[15] R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in multi-agent networked systems. *Proceedings of the IEEE*, 2006. Submitted.

[16] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.

[17] W. Ren, R. W. Beard, and E. M. Atkins. A survey of consensus problems in multi-agent coordination. In *American Control Conference*, pages 1859–1864, Portland, OR, June 2005.

[18] E. P. Ryan. Finite-time stabilization of uncertain nonlinear planar systems. *Dynam. Control*, 1:83–94, 1991.

[19] D. Shevitz and B. Paden. Lyapunov stability theory of nonsmooth systems. *IEEE Transactions on Automatic Control*, 39(9):1910–1914, 1994.

[20] H. Tanner, A. Jadbabaie, and G. J. Pappas. Stable flocking of mobile agents, Part I: Fixed topology. In *IEEE Conf. on Decision and Control*, pages 2010–2015, Maui, HI, December 2003.