Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Achieving coordination tasks in finite time via nonsmooth gradient flows

Jorge Cortés

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 the design of multi-agent coordination algorithms, paying special attention to their scalability properties. Finally, we consider network consensus problems and show how the proposed nonsmooth gradient flows achieve the desired coordination task in finite time.**

I. Introduction

Problem statement: Let $f : \mathbb{R}^d \to \mathbb{R}$, $d \in \mathbb{N}$, be a differentiable function. Consider the gradient dynamical system

$$\dot{x} = -\operatorname{grad}(f)(x).$$

It is well known (see e.g. [1]) that the minima of f are stable equilibria for 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. (A point $x_* \in \mathbb{R}^d$ is a critical point of f if the gradient of f evaluated at x_* vanishes). Gradient dynamical systems are employed to solve problems in a wide range of applications, including optimization, distributed parallel computing, motion planning and control. In robotics, potential field methods are used to autonomously navigate a robot in a cluttered environment. Gradient algorithms enjoy many important 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 dynamical systems

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

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

Jorge Cortés is with the Department of Applied Mathematics and Statistics, Baskin School of Engineering, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA, jcortes@ucsc.edu

Literature review: Guidelines on how to design dynamical systems for optimization purposes, with a special emphasis on gradient systems, are described in [2]. The book [3] 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 are presented in [5], [6], [7]. Finite-time discontinuous feedback stabilizers for a class of planar systems are proposed in [8]. Finite-time stability of continuous autonomous systems is rigorously studied in [9]. The reference [10] 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 a variety of tasks, including cohesiveness [11], [12], [13], consensus [14], and deployment [15], [16]. The distributed algorithms proposed in these papers achieve the desired coordination task asymptotically 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 these flows attain in finite time the set of critical points of the function. To do this, we extend recent results on the stability and convergence properties of general nonsmooth dynamical systems via locally Lipschitz and regular Lyapunov functions. In particular, we develop two novel results involving second-order information about the evolution of the Lyapunov function along solutions of the system to establish finite-time convergence.

We explore the application of the results on nonsmooth gradient flows to the design of multi-agent coordination algorithms. Consider a coordination algorithm defined via the gradient of an aggregate objective function that encodes a desired task. We analyze the algorithms designed via the normalized and signed versions of the gradient, and characterize their scalability properties via the notion of spatially distributed map. We also show how network consensus problems fit nicely into this scheme. We propose two coordination algorithms based on the Laplacian of the communication graph that are guaranteed to achieve consensus in finite time. The normalized gradient descent of the Laplacian potential is not distributed over the communication graph and achieves average-consensus, i.e., consensus at the average of the initial agents' states. The signed gradient descent of the Laplacian potential is distributed over the communication graph and achieves average-max-min-consensus, i.e., consensus at the average of the maximum and the minimum values of the

0-7803-9568-9/05/$20.00 ©2005 IEEE

initial agents' states. Because of length constraints, we refer the interested reader to [17] for the proofs of all the results presented here.

Organization: Section II introduces differential equations with discontinuous right-hand sides and presents various nonsmooth tools for stability analysis. In particular, we develop two novel results involving second-order information and finite-time convergence. Section III introduces the normalized and signed versions of the gradient descent flow of a differentiable function and characterizes their convergence properties. Conditions are given under which these flows converge in finite time. Section IV discusses the application of the results to coordination algorithms for multi-agent systems paying special attention to distributed implementations and network consensus problems. Finally, we gather our conclusions in Section V.

Notation: The set of positive natural numbers is denoted by $\mathbb{N}$. For $d \in \mathbb{N}$, we let $e_1, \dots, e_d$ be the standard orthonormal basis of $\mathbb{R}^d$. For $x \in \mathbb{R}^d$, we denote by $\|x\|_1$ and $\|x\|_2$ the 1-norm and the Euclidean norm of x, respectively. We denote by $v \cdot w$ the inner product of the vectors $v, w \in \mathbb{R}^d$, and by v' the transpose of $v \in \mathbb{R}^d$. For $x \in \mathbb{R}^d$, we let $\operatorname{sgn}(x) = (\operatorname{sgn}(x_1), \dots, \operatorname{sgn}(x_d)) \in \mathbb{R}^d$. We define the vector $\mathbf{1} = (1, \dots, 1)' \in \mathbb{R}^d$. For $S \in \mathbb{R}^d$, we let $\operatorname{co}(S)$ denote its convex closure. We also define $\operatorname{diag}((\mathbb{R}^d)^n) = \{(p, \dots, p) \in (\mathbb{R}^d)^n \mid p \in \mathbb{R}^d\}$ for $n \in \mathbb{N}$. Given a positive semidefinite matrix A, let $H_0(A) \subset \mathbb{R}^d$ denote the eigenspace corresponding to the eigenvalue 0 (if A is positive definite, then we set $H_0(A) = \{0\}$). We denote by $\pi_A : \mathbb{R}^d \to H_0(A)$ the orthogonal projection onto $H_0(A)$. We denote by $\lambda_2(A)$ the smallest non-zero eigenvalue of A, i.e. $\lambda_2(A) = \min\{\lambda \mid \lambda > 0 \text{ and } \lambda \text{ eigenvalue of } A\}$. It is not difficult to see that for any $u \in \mathbb{R}^d$, one has

$$u' A\, u \geq \lambda_2(A)\, \|u - \pi_{H_0(A)}(u)\|_2^2. \tag{1}$$

II. NONSMOOTH STABILITY ANALYSIS

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

A. Differential equations with discontinuous right-hand sides

For differential equations with discontinuous right-hand sides we understand the solutions in terms of differential inclusions following [6]. Let $F : \mathbb{R}^d \to 2^{\mathbb{R}^d}$ be a set-valued map. Consider the differential inclusion

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

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

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

where $X : \mathbb{R}^d \to \mathbb{R}^d$ is measurable and locally essentially bounded [6]. We understand the solution of this equation in the Filippov sense. For $x \in \mathbb{R}^d$, consider the set

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

where μ denotes the usual Lebesgue measure in $\mathbb{R}^d$. A Filippov solution of (3) on an interval $[t_0, t_1] \subset \mathbb{R}$ is defined as a solution of the differential inclusion

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

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. [6]). A set M is *weakly invariant* (respectively *strongly invariant*) for (3) if for each $x_0 \in M$, M contains a maximal solution (respectively all maximal solutions) of (3).

B. Stability analysis via nonsmooth Lyapunov functions

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally Lipschitz function. 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 (cf. [4]) is defined by

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

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

Given a locally Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}$, the *set-valued Lie derivative of* f *with respect to* X at x (cf. [5], [15]) is defined as

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

For each $x \in \mathbb{R}^d$, $\widetilde{\mathcal{L}}_X f(x)$ 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 $\widetilde{\mathcal{L}}_X f(x)$ corresponds to the singleton $\{\mathcal{L}_X f(x)\}$, the usual Lie derivative of f in the direction of X at x. The next result, taken from [5], states that the set-valued Lie derivative allows us to study the evolution of a function along the Filippov solutions.

Theorem 2.1: 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, $\frac{d}{dt}(f(x(t)))$ exists a.e. and $\frac{d}{dt}(f(x(t))) \in \widetilde{\mathcal{L}}_X f(x(t))$ a.e.

In some cases, 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 following result.

Proposition 2.2: 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 a

Lipschitz and regular function. Then $\frac{d^2}{dt^2}(f(x(t)))$ exists a.e. and $\frac{d^2}{dt^2}(f(x(t))) \in \widetilde{\mathcal{L}}_X(\widetilde{\mathcal{L}}_X f)(x(t))$ a.e.

The following result is a generalization of LaSalle principle for differential equations of the form (3) with nonsmooth Lyapunov functions. The formulation is taken from [5], and slightly generalizes the one presented in [7].

Theorem 2.3: (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

$$Z_{X,f} = \left\{ x \in \mathbb{R}^d \mid 0 \in \widetilde{\mathcal{L}}_X f(x) \right\} .$$

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 $\overline{Z}_{X,f} \cap S$. Furthermore, 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 [15].

Proposition 2.4: (Finite-time convergence with first-order information): Under the same assumptions of Theorem 2.3, further assume that there exists a neighborhood U of $Z_{X,f} \cap S$ in S such that $\max \widetilde{\mathcal{L}}_X f < -\epsilon < 0$ 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 $x_0 \in S$ attains $Z_{X,f} \cap S$ in finite time.

Often times, first-order information is inconclusive to assess the finite-time convergence of an specific flow. The next result makes use of second-order information to arrive at a satisfactory answer.

Theorem 2.5: (Finite-time convergence with second-order information): Under the same assumptions of Theorem 2.3, further assume that

(i) the function $x \in \mathbb{R}^d \mapsto \widetilde{\mathcal{L}}_X f(x)$ is single-valued, Lipschitz and regular;
(ii) there exists a neighborhood U of $Z_{X,f} \cap S$ in S such that $\max \widetilde{\mathcal{L}}_X(\widetilde{\mathcal{L}}_X f) > \epsilon > 0$ 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 $x_0 \in S$ attains $Z_{X,f} \cap S$ in finite time.

III. NONSMOOTH GRADIENT FLOWS WITH FINITE-TIME CONVERGENCE

In this section, we formally introduce the normalized and signed gradient dynamical systems associated with a differentiable function. We characterize their general asymptotic convergence properties. Building on the novel results of the previous section, we identify conditions on the differentiable function under which convergence is reached in finite time.

Consider the following dynamical systems on $\mathbb{R}^d$

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

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

Clearly, both differential equations in (6) have discontinuous right-hand sides. Therefore, we understand their solutions in the Filippov sense. The following result describes their associated set-valued maps.

Lemma 3.1: The Filippov set-valued maps associated with the discontinuous vector fields of equations (6a) and (6b) are described by

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

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

Note in particular that $K\Big[\frac{\operatorname{grad}(f)}{\|\operatorname{grad}(f)\|_2}\Big](x) = \frac{\operatorname{grad}(f)(x)}{\|\operatorname{grad}(f)(x)\|_2}$ 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 the vector fields in equations (6a) and (6b).

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

Proposition 3.2: 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) (respectively, for (6b)). Then each solution of equation (6a) (respectively equation (6b)) starting from x_0 asymptotically converges to $\operatorname{Critical}(f)$.

Let us now discuss the finite-time convergence properties of the vector fields (6). Note that Proposition 2.4 cannot be applied to these flows. Indeed, one has

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

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

and $\inf_{x\in U\setminus\operatorname{Critical}(f)\cap S} \|\operatorname{grad}(f)(x)\|_2 = 0$ 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. Therefore, the hypotheses of Proposition 2.4 are not verified by either the flow (6a) or the flow (6b).

Under some additional conditions on the function f, one can establish stronger convergence properties of the solutions of equations (6). We show this in the following result.

Theorem 3.3: 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) (respectively, 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 $x \in V$, the Hessian $\operatorname{Hess}(f)(x)$ is positive definite; or
(ii) for all $x \in V \setminus (\operatorname{Critical}(f) \cap S)$, the Hessian $\operatorname{Hess}(f)(x)$ 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 the eigenvalue 0.

Then each solution of equation (6a) (respectively equation (6b)) starting from x_0 converges in finite time to a critical point of f.

Corollary 3.4: 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) (respectively, for (6b)). Assume that for each $x \in \text{Critical}(f) \cap S$, the Hessian $\text{Hess}(f)(x)$ is positive definite. Then each solution of equation (6a) (respectively equation (6b)) starting from x_0 converges in finite time to a minimum of f.

IV. Applications to coordination algorithms for multi-agent systems

Here we discuss the application of the results on the proposed nonsmooth gradient dynamical systems to the design of multi-agent coordination algorithms. We start by presenting the notion of proximity graphs from computational geometry and of spatially distributed map. Given a coordination algorithm defined via the gradient of an aggregate objective function, these concepts will allow us to characterize the scalability properties of the coordination algorithms designed via the normalized and signed versions of the gradient. We end the section illustrating our results in network consensus problems.

A. Proximity graphs and spatially-distributed maps

We introduce some concepts regarding proximity graphs for point sets in $\mathbb{R}^d$. We assume the reader is familiar with the standard notions of graph theory as defined in [18, Chapter 1]. We begin with some notation. Given a vector space $\mathbb{V}$, let $\mathbb{F}(\mathbb{V})$ be the collection of finite subsets of $\mathbb{V}$. Accordingly, $\mathbb{F}(\mathbb{R}^d)$ is the collection of finite point sets in $\mathbb{R}^d$; elements of $\mathbb{F}(\mathbb{R}^d)$ are of the form $\{p_1, \dots, p_m\} \subset \mathbb{R}^d$, where $p_1, \dots, p_m$ are distinct points in $\mathbb{R}^d$. Let $\mathbb{G}(\mathbb{R}^d)$ be the set of undirected graphs whose vertex set is an element of $\mathbb{F}(\mathbb{R}^d)$. Finally, let $i_{\mathbb{F}} : (\mathbb{R}^d)^n \to \mathbb{F}(\mathbb{R}^d)$ be the natural immersion, i.e., $i_{\mathbb{F}}(P)$ is the point set that contains only the distinct points in $P = (p_1, \dots, p_n) \in (\mathbb{R}^d)^n$. Note that the cardinality of $i_{\mathbb{F}}(p_1, \dots, p_n)$ is in general less than or equal to n.

A *proximity graph function* $\mathcal{G} : (\mathbb{R}^d)^n \to \mathbb{G}(\mathbb{R}^d)$ associates to a tuple $P \in (\mathbb{R}^d)^n$ an undirected graph with vertex set $i_{\mathbb{F}}(\mathcal{P})$ and edge set $\mathcal{E}_{\mathcal{G}}(P)$, where $\mathcal{E}_{\mathcal{G}} : (\mathbb{R}^d)^n \to \mathbb{F}(\mathbb{R}^d \times \mathbb{R}^d)$. In other words, the edge set of a proximity graph depends on the location of its vertices. Examples of proximity graphs include the complete graph, the r-disk graph, the Euclidean Minimum Spanning Three, the Delaunay graph, etc. see [19], [20], [16]. To each proximity graph $\mathcal{G}$, one associates the *set of neighbors map* $\mathcal{N}_{\mathcal{G}} : (\mathbb{R}^d)^n \to (\mathbb{F}(\mathbb{R}^d))^n$, defined by

$$\mathcal{N}_{\mathcal{G},i}(P) = \{p_j \in i_{\mathbb{F}}(P) \mid j \neq i \text{ and } (p_i, p_j) \in \mathcal{E}_{\mathcal{G}}(P)\}.$$

Note that any standard directed graph G with vertex set $\{1, \dots, n\}$ and edge set $E \subset \{1, \dots, n\} \times \{1, \dots, n\}$ can be seen as a proximity graph where, for each $P \in (\mathbb{R}^d)^n$, $(p_i, p_j) \in \mathcal{E}_{\mathcal{G}}(P)$ if and only if $(i, j) \in E$. In this case, $\mathcal{N}_{\mathcal{G},i}(P) = \mathcal{N}_{G,i} = \{j \in \{1, \dots, n\} \mid (i, j) \in E\}$.

Given a set Y and a proximity graph function $\mathcal{G}$, a map $T : (\mathbb{R}^d)^n \to Y^n$ is *spatially distributed over* $\mathcal{G}$ if there exist a map $\widetilde{T} : \mathbb{R}^d \times \mathbb{F}(\mathbb{R}^d) \to Y$, with the property that, for all $(p_1, \dots, p_n) \in (\mathbb{R}^d)^n$ and for all $j \in \{1, \dots, n\}$,

$$T_j(p_1, \dots, p_n) = \widetilde{T}(p_j, \mathcal{N}_{\mathcal{G},j}(p_1, \dots, p_n)),$$

where T_j denotes the jth-component of T. In other words, the jth component of a spatially distributed map at $(p_1, \dots, p_n)$ can be computed with only the knowledge of the vertex p_j and the neighboring vertices in the graph $\mathcal{G}(p_1, \dots, p_n)$.

B. Gradient coordination algorithms

There are a number of gradient-following algorithms proposed in the literature to optimize aggregate objective functions encoding various coordination tasks. Examples include the deployment algorithms in [16], the consensus algorithm in [14] and the cohesiveness algorithms in [11], [12], [13].

The general idea is the following: consider a network composed of n agents with sensing, computing, communication, and motion control capabilities. The state of the ith agent, denoted by $p_i \in \mathbb{R}^d$, might correspond, depending on the specific problem, to the location of the agent in space, or to other physical quantities like attitude, temperature, or voltage. This state p_i evolves according to a first-order continuous dynamics of the form

$$\dot{p}_i(t) = u_i. \tag{7}$$

Here, the control u_i takes values in a bounded subset of $\mathbb{R}^d$. Additionally, the communication topology of the network is described by a proximity graph $\mathcal{G}$. Specifically, the ith agent is capable of transmitting information to the jth agent if and only if $p_j \in \mathcal{N}_{\mathcal{G},p_i}(\mathcal{P})$. Typical proximity graphs employed are the r-disk graph (where two agents are neighbors if they are at a distance at most $r \in \mathbb{R}_+$ from each other) or the visibility graph (where two agents are neighbors if they are visible to each other), see [16].

The last ingredient is an aggregate objective function $\mathcal{H} : (\mathbb{R}^d)^n \to \mathbb{R}$ with two important properties: (i) its critical points correspond to the network configurations where the desired coordination task is achieved, and (ii) its gradient $\text{grad}(\mathcal{H}) : (\mathbb{R}^d)^n \to (\mathbb{R}^d)^n$ is spatially distributed over the proximity graph $\mathcal{G}$. One then sets up the gradient coordination algorithm

$$\dot{p}_i(t) = -\frac{\partial \mathcal{H}}{\partial p_i}(p_1(t), \dots, p_n(t)), \quad i \in \{1, \dots, n\}, \tag{8}$$

which is spatially distributed over $\mathcal{G}$.

Following equations (6), consider the normalized and signed versions of the gradient coordination algorithm (8)

$$\dot{p}_i = -\frac{\frac{\partial \mathcal{H}}{\partial p_i}(p_1(t), \dots, p_n(t))}{\|\frac{\partial \mathcal{H}}{\partial P}(p_1(t), \dots, p_n(t))\|_2}, \tag{9a}$$

$$\dot{p}_i = -\operatorname{sgn}\Big(\frac{\partial \mathcal{H}}{\partial p_i}(p_1(t), \dots, p_n(t))\Big). \tag{9b}$$

Although both vector fields enjoy similar convergence properties (as established by Proposition 3.2), there is a fundamental difference between them, as the next result states.

Proposition 4.1: Let $\mathcal{G}$ be a proximity graph other than the complete graph. Let $\mathcal{H} : (\mathbb{R}^d)^n \to \mathbb{R}$ such that $\text{grad}(\mathcal{H}) : (\mathbb{R}^d)^n \to (\mathbb{R}^d)^n$ is spatially distributed over $\mathcal{G}$. Then

(i) the coordination algorithm in (9a) is not spatially distributed over $\mathcal{G}$;

(ii) the coordination algorithm in (9b) is spatially distributed over $\mathcal{G}$.

Remark 4.2: A different approach to the design of motion coordination algorithms consists of identifying a meaningful local objective function for each agent (i.e., defined with the information provided by its neighbors), whose optimization helps the network achieve the global task, and following its gradient. (e.g. the rendezvous strategies in [21], and the basic interaction laws in [15]). The finite-time convergence properties of the gradient flows (6a) and (6b) can be invaluable in characterizing the asymptotic convergence of the resulting coordination algorithm. In both cases, the algorithms are naturally spatially distributed with respect to the selected proximity graph.

C. Network consensus problems

Here we focus on consensus problems. Let $G = (\{1, \dots, n\}, E)$ be an undirected graph with n vertices. The graph Laplacian matrix L associated with G (see, for instance, [18]) is defined as $L = \Delta - A$, where Δ is the degree matrix and A is the adjacency matrix of the graph. The Laplacian matrix has the following relevant properties: it is symmetric, positive semidefinite and has $\lambda = 0$ as an eigenvalue with eigenvector $\mathbf{1}$. More importantly, the graph G is connected if and only if $\text{rank}(L) = n - 1$, i.e., if the eigenvalue 0 has multiplicity one. This is the reason why the eigenvalue $\lambda_2(L) = \min\{\lambda \mid \lambda > 0 \text{ and } \lambda \text{ eigenvalue of } L\}$ is termed the *algebraic connectivity* of the graph G.

In this setting, the agents' states p_i, $i \in \{1, \dots, n\}$, evolve in $\mathbb{R}$, $p_i \in \mathbb{R}$. The variable p_i does not necessarily refer to physical variables such as spatial coordinates or velocities. Two agents p_i and p_j are said to *agree* if and only if $p_i = p_j$. A meaningful function that quantifies the group disagreement in a network is the so-called *disagreement function* or *Laplacian potential* $\Phi_G : \mathbb{R}^n \to \overline{\mathbb{R}}_+$ associated with G (see [14]), defined by

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

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 the graph G is connected, then all nodes in the graph 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

$$\text{grad}(\Phi_G)(P) = LP\,,$$

which is clearly spatially distributed over the proximity graph induced by G. The gradient coordination algorithm

$$\dot{p}_i = -\frac{\partial \Phi_G}{\partial p_i} = \sum_{j \in \mathcal{N}_{G,i}} (p_j - p_i), \quad i \in \{1, \dots, n\} \tag{10}$$

asymptotically converges to the critical points of Φ_G, i.e., asymptotically achieves consensus. Actually, since the system is linear, the convergence is exponential. 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 (10) 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*).

Now, consider the discontinuous differential equations corresponding to (6), for $i \in \{1, \dots, n\}$,

$$\dot{p}_i = \frac{\sum_{j \in \mathcal{N}_{G,i}} (p_j - p_i)}{\|LP\|_2}, \tag{11a}$$

$$\dot{p}_i = \text{sgn}\big(\sum_{j \in \mathcal{N}_{G,i}} (p_j - p_i) \big). \tag{11b}$$

Before analyzing the convergence properties of these flows, let us identify a conserved quantity for each one of them.

Proposition 4.3: 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 (11a) and g_2 is constant along the solutions of (11b).

The following theorem completely characterizes the asymptotic convergence properties of the flows in (11).

Theorem 4.4: Let $G = (\{1, \dots, n\}, E)$ be a connected undirected graph. Then, the flows in (11) 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 (11a) 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);

(ii) the solutions of (11b) 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).

Remark 4.5: An interesting observation regarding the flow (11b) is the following. From Theorem 4.4, the network achieves average-max-min-consensus. From its evolution, one can deduce that the convergence time is given by

$$\frac{1}{2}\big(\max_{i \in \{1, \dots, n\}} \{(p_i)_0\} - \min_{i \in \{1, \dots, n\}} \{(p_i)_0\} \big).$$

In particular, if the network agents had the capability to decide exactly when convergence has been achieved (for instance, by running in parallel another consensus algorithm), then this information together with the consensus value will serve them to compute both the values of $\max_{i \in \{1, \dots, n\}} \{(p_i)_0\}$ and $\min_{i \in \{1, \dots, n\}} \{(p_i)_0\}$. •

Remark 4.6: It is also possible to consider networks with switching communication topologies and establish a result similar to Theorem 4.4 (see [17] for a detailed discussion).•

Fig. 1 illustrates the evolution of the differential equations (10), (11a) and (11b). As stated in Theorem 4.4, the agents evolving under (11a) achieve average-consensus in finite time, and the agents evolving under (11b) achieve average-max-min-consensus in finite time.

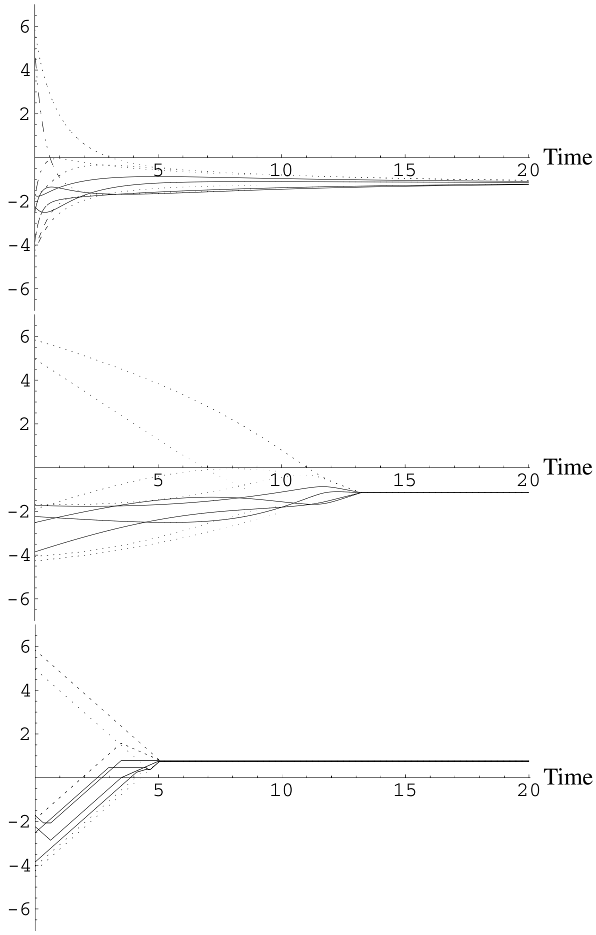

Fig. 1. From top to bottom, evolution of (10), (11a) and (11b) for 10 agents starting from a randomly generated initial configuration with $p_i \in [-7, 7]$, $i \in \{1, \ldots, 10\}$. The graph $G = (\{1, \ldots, 10\}, E)$ has edge set $E = \{(1, 4), (1, 10), (2, 10), (3, 6), (3, 9), (4, 8), (5, 6), (5, 9), (7, 10), (8, 9)\}$. The algebraic connectivity of G is $\lambda_2(L) = 0.12$.

V. Conclusions

We have introduced the normalized and signed versions of the gradient descent flow of a differentiable function. We have characterized the general asymptotic convergence properties of these nonsmooth gradient flows, and identified suitable conditions on the differentiable function that guarantee that convergence to the critical points is achieved in finite time. We have discussed the application of these results to gradient coordination algorithms for multi-agent systems, and, in particular, to consensus problems.

Future work will be devoted to explore (i) the development of tight upper bounds on the (finite) convergence time of the proposed nonsmooth flows. These results promise to be useful in assessing the (time) complexity of distributed coordination algorithms; (ii) the application of the results to distributed sensor fusion algorithms based on consensus (e.g. [22], [23]), coordination problems such as formation control, deployment and rendezvous, and other problems where gradient systems play an important role; (iii) the identification of more nonsmooth distributed algorithms based on gradient information with similar convergence properties.

Acknowledgments

This research was supported in part by faculty research funds granted by UCSC, and NASA University Aligned Research Program Award TO.030.MM.D.

References

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

[2] U. Helmke and J. Moore, *Optimization and Dynamical Systems*. New York: Springer Verlag, 1994.

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

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

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

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

[7] D. Shevitz and B. Paden, "Lyapunov stability theory of nonsmooth systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 9, pp. 1910–1914, 1994.

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

[9] S. P. Bhat and D. S. Bernstein, "Finite-time stability of continuous autonomous systems," *SIAM Journal on Control and Optimization*, vol. 38, no. 3, pp. 751–766, 2000.

[10] 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*, vol. 33, pp. 804–833, 1995.

[11] V. Gazi and K. M. Passino, "Stability analysis of swarms," *IEEE Transactions on Automatic Control*, vol. 48, no. 4, pp. 692–697, 2003.

[12] 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*, vol. 49, no. 8, pp. 1292–1302, 2004.

[13] H. Tanner, A. Jadbabaie, and G. J. Pappas, "Flocking in fixed and switching networks," *IEEE Transactions on Automatic Control*, 2005. Submitted.

[14] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[15] J. Cortés and F. Bullo, "Coordination and geometric optimization via distributed dynamical systems," *SIAM Journal on Control and Optimization*, 2005. To appear.

[16] J. Cortés, S. Martínez, and F. Bullo, "Spatially-distributed coverage optimization and control with limited-range interactions," *ESAIM. Control, Optimisation & Calculus of Variations*, 2005. To appear.

[17] J. Cortés, "Finite-time convergent gradient flows with applications to motion coordination," *Systems & Control Letters*, Sept. 2005. Submitted.

[18] R. Diestel, *Graph Theory*, vol. 173 of *Graduate Texts in Mathematics*. New York: Springer Verlag, second ed., 2000.

[19] M. de Berg, M. van Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*. New York: Springer Verlag, 1997.

[20] J. W. Jaromczyk and G. T. Toussaint, "Relative neighborhood graphs and their relatives," *Proceedings of the IEEE*, vol. 80, no. 9, pp. 1502–1517, 1992.

[21] H. Ando, Y. Oasa, I. Suzuki, and M. Yamashita, "Distributed memoryless point convergence algorithm for mobile robots with limited visibility," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 5, pp. 818–828, 1999.

[22] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, "Distributed sensor fusion using dynamic consensus," in *IFAC World Congress*, (Prague, CZ), July 2005. Electronic proceedings.

[23] L. Xiao, S. Boyd, and S. Lall, "A scheme for asynchronous distributed sensor fusion based on average consensus," in *Fourth International Conference on Information Processing in Sensor Networks (IPSN'05)*, (Los Angeles, CA), 2005. To appear.