Nonsmooth Neural Network for Convex Time-Dependent Constraint Satisfaction Problems

Mauro Di Marco, Mauro Forti, Paolo Nistri, Luca Pancioni

***Abstract*—The paper introduces a nonsmooth (NS) neural network which is able to operate in a time-dependent (TD) context and is potentially useful for solving some classes of NS-TD problems. The proposed network is named NTN and is an extension to a TD setting of a previous NS neural network for programming problems. Suppose $C(t)$, $t \geq 0$, is a nonempty TD convex feasibility set defined by TD inequality constraints. The constraints are in general NS (nondifferentiable) functions of the state variables and time. NTN is described by the subdifferential with respect to the state variables of a NS-TD barrier function and a vector field corresponding to the unconstrained dynamics. The paper shows that for suitable values of the penalty parameter the NTN dynamics displays two main phases. In the first phase any solution of NTN not starting in $C(0)$ at $t = 0$ is able to reach the moving set $C(\cdot)$ in finite time t_h, whereas in the second phase the solution tracks the moving set, i.e., it stays within $C(t)$ for all subsequent times $t \geq t_h$. NTN is thus able to find an exact feasible solution in finite time and also to provide an exact feasible solution for subsequent times. This new and peculiar dynamics displayed by NTN is potentially useful for addressing some significant TD signal processing tasks. As an illustration the paper discusses a number of examples where NTN is applied to the solution of NS-TD convex feasibility problems.**

***Index Terms*—Nonsmooth neural networks; time-dependent constraints; finite-time convergence; convex functions; subdifferential.**

I. Introduction

The analog neural network approach for solving in real time optimization problems with constraints has been originally devised by Tank and Hopfield [1] and further elaborated by Kennedy and Chua [2]. The idea is to use a dynamic neural network that simulates both the objective function and constraints, and to exploit the analog and parallel processing capabilities of the network for computing in real time the optimal solution of the problem being modeled. One important aspect is that in significant applications, as linear and quadratic programming problems, the obtained neural networks can be effectively implemented as VLSI electronic circuits [2]. Since the seminal papers [1], [2], such an approach has been the subject of an intense investigation, and several relevant contributions have been put forward demonstrating its effectiveness in all those problems in the field of signal processing, pattern recognition and control, where finding in real-time optimal feasible solutions is mandatory [1]–[12].

During recent years significant attention has been devoted to certain classes of nonsmooth (NS) dynamic neural networks for programming problems. In particular, [13] has introduced a generalized nonlinear programming circuit (G-NPC) which extends to a NS setting the programming circuit in [2]. G-NPC is able to optimize NS (nondifferentiable) objective functions by means of an exact penalty function method based on NS barrier functions implemented by constraint neuron with high-gain nonlinearities. One key advantage of the NS approach in G-NPC, with respect to that based on smooth circuits in [2], is the presence of sliding modes that enable convergence in finite time toward the feasibility set defined by the constraints, and also convergence in finite time to the optimal solution set for some classes of problems, such as linear programming problems. In addition to G-NPC, several other classes of NS neural networks have been introduced and theoretically analyzed in the literature. They differ in terms of architecture and circuit implementation (use of one or more layers of neurons and various classes of NS nonlinearities), the NS method used (penalty function, projection methods, or smooth approximations), and the hypotheses on the constraints (convex or nonconvex constraints or some generalizations of the notion of convex constraints) [10], [14]–[26]. The study of NS dynamic neural networks for programming problems continues to be a timely and widely investigated topic as witnessed by the several recent publications in the neural network related literature [14], [15], [17]–[19], [24], [26].

G-NPC, and all previous NS neural network models, are actually designed by assuming that the cost functions and constraints do not depend on time. However, several engineering problems concern systems or devices operating in a dynamic environment, so that they are better modeled when accounting for the dependence upon time of the parameters, constraints, and objective functions involved. TD constraints arise in several diverse areas as scheduling, vehicle routing in time-varying environments, control of nonlinear systems with time-varying output constraints, robot planning in dynamic environments, image processing and medical imaging, econometrics, forecasting, fast encryption algorithms, broadband pricing, speech recognition [27]–[38]. Sometimes also TD objective functions are considered as in model predictive control with time-varying terminal cost [39]. It is thus of interest to investigate whether some NS neural network architectures designed for static problems can be suitably modified and extended to work in a TD context and to see if there are significant problems involving TD parameters, constraints, or cost functions, that are amenable to an effective solution via such extended neural networks.

In this paper we introduce a new NS neural network, derived

Manuscript received —; revised —.

The authors are with the Department of Information Engineering and Mathematical Sciences, University of Siena, 53100 Siena, Italy (e-mail: dimarco@dii.unisi.it, forti@dii.unisi.it, pnistri@dii.unisi.it, pancioni@dii.unisi.it).

from G-NPC [13], that is able to operate in a TD context and is useful for solving some classes of NS-TD problems. We will refer to the neural network as NTN. Suppose we are given a nonempty TD convex set $C(t) \subset \mathbb{R}^q$ (feasibility set), for $t \geq 0$, defined by TD inequality constraints. The constraints are allowed to be NS, i.e., nondifferentiable with respect to the state variables $x \in \mathbb{R}^q$ and with respect to time t. The proposed neural network is described by the subdifferential with respect to x of a NS-TD barrier function constructed with the constraints and an additional NS-TD vector field describing the unconstrained dynamics. Under suitable assumptions on the penalty parameter it is shown that the NTN dynamics is characterized by two salient phases. In the first phase, any solution $x(\cdot)$ of NTN with initial condition $x(0) \notin C(0)$ reaches the *moving set* $C(\cdot)$ in *finite time* t_h, i.e., $x(t_h) \in C(t_h)$, where t_h can be uniformly estimated on the basis of the network parameters. In the second phase the network is able to track the moving set, i.e., we have $x(t) \in C(t)$ for any $t \geq t_h$. In other words, by exploiting an exact NS-TD penalty function method, and the sliding modes of the NS dynamics, NTN is able to find an exact feasible solution in finite time t_h and yield an exact feasible solution for all subsequent times $t \geq t_h$, even for TD constraints. In the paper we provide an interpretation of the tracking phase as a Moreau sweeping process [40], and we also show that NTN admits a simple electronic implementation along the lines of previous neural networks for programming problems.

To the authors' knowledge, the present paper is the first investigation on NS neural networks for solving some NS-TD problems. As an illustration of the potential applications of the new and peculiar dynamics displayed by NTN, in the paper we work out a number of examples where NTN is used in the solution of NS-TD convex feasibility problems. Other potential applications of NTN concern the real-time tracking of moving objects and the modeling of systems displaying sweeping processes.

The paper is organized as follows. After providing the needed preliminaries in Section I-A, in Section II we present the neural network model NTN studied in the paper. Section III analyzes the main aspects of the dynamical behavior of NTN, whereas Section IV discusses the dynamics within the framework of the sweeping processes. Then, Section V presents some applications of NTN to convex feasibility problems. Finally, Section VI collects some concluding remarks.

Notation: By $\mathbb{R}$ we denote the set of real numbers, by $\mathbb{R}^+$ the set of nonnegative real numbers and by $\mathbb{R}^q$ the real q-space. Given $x, y \in \mathbb{R}^q$, $\langle x, y\rangle = \sum_{i=1}^q x_i y_i$ is the scalar product of x and y. Moreover, $\|x\| = (\sum_{i=1}^q x_i^2)^{1/2}$ is the Euclidean norm of $x \in \mathbb{R}^q$, whereas $\|x\|_\infty = \max_{i=1,2,\ldots,q} |x_i|$ is the infinity-norm of $x \in \mathbb{R}^q$. Given $r > 0$ and $x \in \mathbb{R}^q$, $B(x,r) = \{y \in \mathbb{R}^n : \|y - x\| < r\}$ is the open ball with radius r centered at x. If $Q \subset \mathbb{R}^q$, $\operatorname{int} Q$ denotes the interior of Q whereas $\operatorname{bd} Q$ is the boundary of Q. The distance of $x \in \mathbb{R}^q$ from $Q \subset \mathbb{R}^q$ is defined as $\operatorname{dist}(x, Q) = \inf_{y \in Q} \|x - y\|$.

A. Preliminaries

1) Upper semicontinuous maps: Let $(t, x) \to F(t, x)$ be a multivalued map from $\mathbb{R}^+ \times \mathbb{R}^q$ into the subsets of $\mathbb{R}^q$ such that $F(t, x)$ is nonempty for any $(t, x) \in \mathbb{R}^+ \times \mathbb{R}^q$. F is said to be upper semicontinuous (u.s.c.) at $(t, x) \in \mathbb{R}^+ \times \mathbb{R}^q$ if for any open set V in $\mathbb{R}^q$ containing $F(t, x)$ there exists a neighborhood U in $\mathbb{R}^+ \times \mathbb{R}^q$ of (t, x) such that $F(U) \subset V$ [41]. When F has closed values, and it is locally bounded, i.e., it is bounded in a neighborhood of each point $(t, x) \in \mathbb{R}^+ \times \mathbb{R}^q$, it can be shown that F is u.s.c. on $\mathbb{R}^q$ if and only if its graph $\{((t, x), y) \in (\mathbb{R}^+ \times \mathbb{R}^q) \times \mathbb{R}^q : y \in F(t, x)\}$ is a closed set.

2) Subdifferential of a convex function: We recall some notions of convex analysis which are needed in the paper. Let $g : \mathbb{R}^q \to \mathbb{R}$ be a convex function. Then g is continuous and locally Lipschitz but not necessarily differentiable. This notwithstanding we can introduce the notion of the subdifferential of g as a generalization of the ordinary differentiation as follows [42, Def. 1.2.1].

Definition 1: The subdifferential $\partial g(x)$ of g at $x \in \mathbb{R}^q$ is the set of vectors $z \in \mathbb{R}^q$ satisfying

$$g(y) \geq g(x) + \langle z, y - x\rangle \quad \forall y \in \mathbb{R}^q.$$

A vector $z \in \partial g(x)$ is called a subgradient of g at x. The map $x \to \partial g(x)$ has nonempty compact convex values, has a closed graph and is locally bounded [42, VI.6.2]. Thus $x \to \partial g(x)$ sends compact sets of $\mathbb{R}^q$ into compact sets of $\mathbb{R}^q$. Moreover, if g is differentiable at x in the ordinary sense, then $\partial g(x) = \{\nabla g(x)\}$ is a singleton.

Throughout the paper we will use the following rules of calculus for the subdifferential of convex functions.

i) Positive combination of convex functions [42, Th. 4.1.1]. Let $g_1, g_2, \ldots, g_m$ be m convex functions from $\mathbb{R}^q$ to $\mathbb{R}$ and $\lambda_1, \lambda_2, \ldots, \lambda_m$ be positive scalars. Then

$$\partial\left(\sum_{i=1}^m \lambda_i g_i\right)(x) = \sum_{i=1}^m \lambda_i \partial g_i(x)$$

for all $x \in \mathbb{R}^q$.

ii) Maximum of convex functions [42, Cor. 4.3.2]. Let $g_1, g_2, \ldots, g_m$ be m convex functions from $\mathbb{R}^q$ to $\mathbb{R}$ and define

$$g(x) = \max\{g_1(x), g_2(x), \ldots, g_m(x)\}$$

for any $x \in \mathbb{R}^q$. If we denote by $I(x) = \{i = 1, 2, \ldots, m : g_i(x) = g(x)\}$ the active index set, we have

$$\partial g(x) = \operatorname{co}\{\cup \partial g_i(x) : i \in I(x)\}$$

where $\operatorname{co}(A)$ denotes the convex hull of the set A.

iii) Chain Rule (see, e.g., [43, Prop. 1]). Suppose that $V : \mathbb{R}^q \to \mathbb{R}$ is a convex function. Also suppose that $x : \mathbb{R}^+ \to \mathbb{R}^q$ is absolutely continuous on any compact interval of $\mathbb{R}^+$. Then, functions $\tau \to x(\tau)$ and $\tau \to V(x(\tau))$ are differentiable for almost every (a.e.) $\tau \geq 0$ and we have

$$\frac{d}{d\tau} V(x(\tau)) = \langle z, \dot{x}(\tau)\rangle \quad \forall z \in \partial V(x(\tau))$$

where the dot denotes the derivative with respect to time.

3) Tangent and normal cones: If $C \subset \mathbb{R}^q$ is a nonempty closed convex set, the tangent cone to C at a point $x \in C$ is defined as [44]

$$T_C(x) = \{v \in \mathbb{R}^n : \liminf_{\rho \to 0^+} \frac{\operatorname{dist}(x + \rho v, C)}{\rho} = 0\}$$

whereas the normal cone to C at $x \in C$ is

$$N_C(x) = \{u \in \mathbb{R}^n : \langle u, v \rangle \leq 0, \forall v \in T_C(x)\}.$$

It can be shown that, for any $x \in \mathbb{R}^q$, $T_C(x)$ and $N_C(x)$ are nonempty closed convex cones in $\mathbb{R}^q$. If we have, in particular, $x \in \mathrm{int}C$, then $T_C(x) = \mathbb{R}^q$ and $N_C(x) = \{0\}$.

4) Hausdorff distance between sets: Let $Q_1, Q_2 \subset \mathbb{R}^q$ be nonempty closed sets. The Hausdorff distance between Q_1 and Q_2 is defined as [45]

$$\mathrm{dist}_{\mathrm{H}}(Q_1, Q_2) = \max\{\sup_{x \in Q_2} \mathrm{dist}(x, Q_1), \sup_{x \in Q_1} \mathrm{dist}(x, Q_2)\}.$$

The Hausdorff distance induces a metric on the nonempty closed sets of $\mathbb{R}^q$. If $Q_1, Q_2 \subset \mathbb{R}^q$ are nonempty compact convex sets, then $\mathrm{dist}_{\mathrm{H}}(Q_1, Q_2) = \mathrm{dist}_{\mathrm{H}}(\mathrm{bd}Q_1, \mathrm{bd}Q_2)$.

II. Neural Network Model

We consider for $t \geq 0$ a TD convex set $C(t) \subset \mathbb{R}^q$ (feasibility set), which is defined by TD inequality constraints as follows. Given continuous TD functions $f_j(t, x) : \mathbb{R}^+ \times \mathbb{R}^q \to \mathbb{R}$, $j = 1, 2, \ldots, p$, we let for any $t \geq 0$

$$C(t) = \{x \in \mathbb{R}^q : f_j(t, x) \geq 0, j = 1, 2, \ldots, p\}.$$

We assume henceforth that $x \to -f_j(t, x)$ is *convex* for any $t \geq 0$ and $j = 1, 2, \ldots, p$. Furthermore, for any $r > 0$ there exist $L_j(r) > 0$ such that $|f_j(t, x_1) - f_j(s, x_2)| \leq L_j(r)(|t - s| + \|x_1 - x_2\|)$ for any $t, s \geq 0$, any $x_1, x_2 \in \mathbb{R}^q$ such that $x_1, x_2 \in B(0, r)$, and $j = 1, 2, \ldots, p$.

Note that we allow functions $(t, x) \to f_j(t, x)$ to be NS, i.e., nondifferentiable with respect to the state-variables x_i, $i = 1, 2, \ldots, q$, or with respect to time, t.

Assumption 1: i) We have $\mathrm{int}C(t) \neq \emptyset$ for any $t \geq 0$;

ii) there exists $R > 0$ such that we have $C(t) \subset B(0, R)$ for any $t \geq 0$;

iii) the multivalued map $t \to C(t)$ is Lipschitz continuous in the Hausdorff metric, i.e., there exists $L_C > 0$ such that

$$\mathrm{dist}_{\mathrm{H}}(C(t), C(s)) \leq L_C|t - s|$$

for any $t, s \geq 0$.

Due to the Lipschitz continuity of the multivalued map $t \to C(t)$, by [41, Th. 1, p. 77] there exists a selection $\xi(t) \in \mathrm{int}C(t)$, $t \geq 0$, which is Lipschitz with constant $L_\xi > 0$, i.e., $\|\xi(t) - \xi(s)\| \leq L_\xi|t - s|$ for any $t, s \geq 0$. For the selection $t \to \xi(t)$ we assume the following.

Assumption 2: There exists $\psi > 0$ such that we have $f_j(t, \xi(t)) \geq \psi$ for any $t \geq 0$ and $j = 1, 2, \ldots, p$.

Under Assumption 1 the set $C(t)$ is nonempty, compact and convex for any $t \geq 0$. The compactness assumption is reasonable since in practice the neural network dynamics should be confined within a set bounded by the saturation levels of the electronic devices (e.g., operational amplifiers) implementing the neurons. Assumption 1-iii) on Lipschitz continuity of $C(\cdot)$ is weak and can be easily checked in practice for sets of interest in the engineering applications, see, e.g., the examples in Section V. Also Assumption 2 is quite a weak hypothesis which is easily satisfied in practice, see again Section V.

Our goal is to design a neural network such that any solution not starting in $C(0)$ at $t = 0$ reaches the moving set $C(\cdot)$ in finite time and stays within $C(\cdot)$ thereafter. To this end we define a NS-TD barrier function $\mathcal{W}(t, x) : \mathbb{R}^+ \times \mathbb{R}^q \to \mathbb{R}$ as

$$\mathcal{W}(t, x) = \sum_{j=1}^{p} \int_0^{f_j(t,x)} d(\rho) d\rho$$

where $\sigma > 0$ and

$$d(\rho) = \begin{cases} 0, & \rho > 0 \\ -\sigma, & \rho < 0 \end{cases}$$

is a discontinuous function. We have $\mathcal{W}(t, x) = \sigma W(t, x)$, where

$$W(t, x) = \sum_{j=1}^{p} \max\{0, -f_j(t, x)\} \tag{1}$$

and $\sigma > 0$ is a threshold (*penalty parameter*) to be chosen later. Since $x \to -f_j(t, x)$, $j = 1, 2, \ldots, p$, are convex functions, also $x \to \max\{0, -f_j(t, x)\}$, $j = 1, 2, \ldots, p$, and $x \to W(t, x)$ are convex for any $t \geq 0$.

The proposed neural network, which is named NTN, is described by the differential inclusion

$$\dot{x} \in F(t, x) - \sigma \partial_x W(t, x) \tag{2}$$

for a.e. $t \geq 0$, where $\partial_x W(t, x)$ is the subdifferential of the convex function $x \to W(t, x)$. As in any penalty function method the term $-\sigma \partial_x W(t, x)$ is introduced to enforce constraint satisfaction.

For the sake of generality in model (2) we allow for the presence of a vector field $F(t, x)$ describing the unconstrained dynamics of NTN. $F(t, x)$ can account for example of modeling errors and other inaccuracies, whereas in the applications to sweeping processes it can model external forces acting on the dynamical system (cf. Section IV). By means of $F(t, x)$ we can in practice address a more general problem where we wish that any solution of NTN reaches and tracks a moving set in a robust fashion, i.e., even in the presence of modeling errors or an unconstrained dynamics.

Assumption 3: The set-valued map $F : \mathbb{R}^+ \times \mathbb{R}^q \to \mathbb{R}^q$ is u.s.c. with nonempty compact convex values. Moreover, there exists $\kappa > 0$ such that we have $F(t, x) \subseteq \kappa(1 + \|x\|)B(0, 1)$ for any $(t, x) \in \mathbb{R}^+ \times \mathbb{R}^q$.

Considering that, for any fixed $t \geq 0$, $x \to \max\{0, -f_j(t, x)\}$, $j = 1, 2, \ldots, p$, are convex functions, we have $\partial_x \sum_{j=1}^{p} \max\{0, -f_j(t, x)\} = \sum_{j=1}^{p} \partial_x \max\{0, -f_j(t, x)\}$, see point i) of Section I-A2. Then, by arguing as in the proof of Property 3 in [13], and accounting for the rule in point ii) of Section I-A2, the NTN equations can be rewritten as

$$\dot{x}_i \in F_i(t, x) - \sum_{j=1}^{p} \hat{d}(f_j(t, x)) \partial_{x,i} f_j(t, x)$$

for a.e. $t \geq 0$, and for any $i = 1, 2, \ldots, q$, where

$$\hat{d}(\rho) = \begin{cases} 0, & \rho > 0 \\ [-\sigma, 0], & \rho = 0 \\ -\sigma, & \rho < 0 \end{cases}. \tag{3}$$

Since NTN is defined by a differential inclusion (2), we need to define what is meant by a solution to a Cauchy problem associated with (2). Given $x_0 \in \mathbb{R}^q$, and $\tau > 0$, by a (local) solution on $[0, \tau]$ of (2), with initial condition x_0, we mean an absolutely continuous function $x(\cdot)$ on $[0, \tau]$ such that $x(0) = x_0$ and $\dot{x}(t) \in F(t, x(t)) - \sigma \partial_x W(t, x(t))$ for a.e. $t \in [0, \tau]$. We say that $x(t)$, $t \geq 0$, is a (global) solution of (2) for $t \geq 0$ if $x(t)$, $t \in [0, \tau]$, is a (local) solution of (2) for any $\tau > 0$.

Property 1: If Assumption 3 is satisfied, then for any $x_0 \in \mathbb{R}^q$ there exists at least a local solution of (2) with initial condition x_0.

Proof: For any $t \geq 0$, $x \to W(t, x)$ is convex, and thus $x \to \partial_x W(t, x)$ has nonempty compact convex values (Section I-A2). Moreover, it can be verified that $(t, x) \to \partial_x W(t, x)$ is locally bounded with closed graph, hence it is u.s.c. (Section I-A1). In conclusion, under Assumption 3, for any $\sigma > 0$ the right-hand side of (2) is u.s.c. with nonempty compact convex values as the sum of two maps enjoying this property. The result then follows from [41, p. 98, Th. 3] since the map $(t, x) \to m(F(t, x) - \sigma \partial_x W(t, x))$, where $m(K)$ is the element of K with the smallest norm, is locally compact. ■

Finally, we note that NTN is an extension to a TD context of G-NPC [13]. In fact, in the special case where F and the inequality constraints f_j do not depend on time, and we have $F(x) = -\partial \phi(x)$, where $\phi : \mathbb{R}^q \to \mathbb{R}$ is regular and $\partial \phi(x)$ denotes the Clarke's gradient of ϕ at x, the differential inclusion (2) coincides with that describing G-NPC.

III. Dynamical analysis

In Section III-A we study the boundedness, prolongability, and uniqueness of solutions of NTN with respect to the initial conditions. Then we show that, by suitably choosing the penalty parameter σ, NTN displays two main phases of the dynamics: given a moving set $C(t)$, $t \geq 0$, any solution $x(t)$ of NTN with $x(0) \in B(0, R) \backslash C(0)$ is such that there exists $0 < t_h < +\infty$ for which $x(t_h) \in C(t_h)$, i.e., the solution reaches $C(\cdot)$ in finite time t_h (Section III-B). Moreover, we have $x(t) \in C(t)$ for any $t \geq t_h$, i.e., the solution tracks the moving set $C(\cdot)$ for $t \geq t_h$ (Section III-C). Some remarks on these results are given in Section III-D.

A. Boundedness of solutions

We start by establishing a useful preliminary result.

Lemma 1: Suppose that Assumpions 1-2 are satisfied. Then, for any $t \geq 0$, and any $x \in \mathbb{R}^q$ such that $x \notin C(t)$, we have

$$\langle x - \xi(t), z \rangle \geq \psi + W(t, x) > 0 \qquad \forall z \in \partial_x W(t, x)$$

where the selection $t \to \xi(t)$ and constant $\psi > 0$ are defined in Assumption 2.

Proof: If $g(x) : \mathbb{R}^q \to \mathbb{R}$ is a convex function then $\partial g(x)$ is the set of vectors $z \in \mathbb{R}^q$ such that $g(x+u) - g(x) \geq \langle u, z \rangle$ for all $u \in \mathbb{R}^q$ (Definition 1). Suppose we have $g(\xi_\rho) < 0$ and $g(x) > 0$ for some $\xi_\rho, x \in \mathbb{R}^q$. By choosing $u = \xi_\rho - x$ we obtain $0 > g(\xi_\rho) > g(\xi_\rho) - g(x) \geq \langle \xi_\rho - x, z \rangle$, i.e.,

$$\langle x - \xi_\rho, z \rangle \geq g(x) - g(\xi_\rho) > 0. \tag{4}$$

For any $(t, x) \in \mathbb{R}^+ \times \mathbb{R}^q$ we have $W(t, x) = \sum_{j=1}^p \max\{0, -f_j(t, x)\} = \sum_{j \in I^-(t,x)} |f_j(t, x)|$, where $I^-(t, x) = \{j \in \{1, 2, \ldots, p\} : f_j(t, x) < 0\} \neq \emptyset$ since by assumption $x \notin C(t)$. If $z \in \partial_x W(t, x)$, then $z = \sum_{j \in I^-(t,x)} z_j$ where $z_j \in -\partial_x f_j(t, x)$, $j = 1, 2, \ldots, p$. Functions $x \to -f_j(t, x)$ are convex and by Assumption 2 we have $f_j(t, \xi(t)) > \psi$ for any $j = 1, 2, \ldots, p$, whereas $f_j(t, x) < 0$ for $j \in I^-(t, x)$. Then, by (4) with $g(x) \to -f_j(t, x)$ and $\xi_\rho \to \xi(t)$, if $z \in \partial_x W(t, x)$, we obtain

$$\begin{aligned}
\langle x - \xi(t), z \rangle &= \sum_{j \in I^-(t,x)} \langle x - \xi(t), z_j \rangle \\
&\geq \sum_{j \in I^-(t,x)} (-f_j(t, x) + f_j(t, \xi(t))) \\
&\geq \sum_{j \in I^-(t,x)} |f_j(t, x)| + \psi = W(t, x) + \psi.
\end{aligned}$$

■

To study boundedness of solutions of NTN we find it useful to use the squared distance of $x(t)$ from the selection $\xi(t)$

$$\Delta(t) \doteq \frac{1}{2}(\|x(t) - \xi(t)\|)^2, \quad t \geq 0 \tag{5}$$

as a Lyapunov function for NTN.

Lemma 2: Suppose that Assumpions 1-3 are satisfied. Let $x(t)$, $t \in [0, \tau]$, where $\tau > 0$, be a (local) solution of (2) such that $x(0) \notin C(0)$. Then, for a.e. $t \in [0, \tau]$ such that $x(t) \notin C(t)$ we have

$$\dot{\Delta}(t) \leq -\sigma(\psi + W(t, x(t))) + \|x(t) - \xi(t)\|(\|\gamma(t)\| + L_\xi)$$

where $\gamma(t) \in F(t, x(t))$ and L_ξ is the Lipschitz constant of the selection $t \to \xi(t)$.

Proof: The solution $t \to x(t)$ is absolutely continuous on $[0, \tau]$. Moreover, since $t \to \xi(t)$ is Lipschitz, by Rademacher theorem it is differentiable for a.e. $t \in [0, \tau]$. Then, for a.e. $t \in [0, \tau]$ also $t \to \Delta(t)$ is differentiable and we have $\dot{\Delta}(t) = \langle x(t) - \xi(t), \dot{x}(t) - \dot{\xi}(t) \rangle$. Since $\dot{x}(t) \in F(t, x(t)) - \sigma \partial_x W(t, x(t))$, there exist $\gamma(t) \in F(t, x(t))$ and $z(t) \in \partial_x W(t, x(t))$ such that for a.e. $t \in [0, \tau]$ we have $\dot{x}(t) = \gamma(t) - \sigma z(t)$. Then, we obtain

$$\begin{aligned}
\dot{\Delta}(t) &= \langle x(t) - \xi(t), \gamma(t) - \sigma z(t) \rangle - \langle x(t) - \xi(t), \dot{\xi}(t) \rangle \\
&\leq -\langle x(t) - \xi(t), \sigma z(t) \rangle \\
&\quad + \|x(t) - \xi(t)\|(\|\gamma(t)\| + \|\dot{\xi}(t)\|) \\
&\leq -\sigma(\psi + W(t, x(t))) + \|x(t) - \xi(t)\|(\|\gamma(t)\| + L_\xi)
\end{aligned}$$

where in the last inequality we considered that, by Lemma 1, $\langle x(t) - \xi(t), z(t) \rangle \geq \psi + W(t, x(t))$ when $x(t) \notin C(t)$. ■

If Assumption 3 holds, we have $F(t, B(0, r)) \subset \kappa(1 + r)B(0, 1)$ for any $t \geq 0$ and $r > 0$. Hence

$$M_F(r) \doteq \sup\{\|F(t, x)\|, (t, x) \in \mathbb{R}^+ \times B(0, r)\} < +\infty$$

for any $r > 0$, where $\|F(t, x)\| = \sup\{\|y\| : y \in F(t, x)\}$.

Proposition 1: Suppose Assumpions 1-3 are satisfied and

$$\sigma > \sigma_a = \frac{2R}{\psi}(M_F(3R) + L_\xi) \tag{6}$$

where $R > 0$ is given in Assumption 1. Then, any solution $x(t)$ of NTN such that $x(0) \in B(0, R)$ is bounded and hence defined for all $t \geq 0$. In particular, we have $x(t) \in B(0, 3R)$ and $\|x(t) - \xi(t)\| < 2R$ for any $t \geq 0$. Moreover, for a.e. $t \geq 0$ such that $x(t) \notin C(t)$ we have

$$\dot{\Delta}(t) < -2Rh^2 < 0 \tag{7}$$

where

$$h^2 = \frac{\psi}{2R}(\sigma - \sigma_a) > 0. \tag{8}$$

Proof: Suppose I_x is the maximal interval of existence of the solution $t \to x(t)$ and let $\tau > 0$ be such that $[0, \tau] \subset I_x$. We want to show that $\|x(t) - \xi(t)\| < 2R$, and so, by Assumption 1, $\|x(t)\| < 3R$, for any $t \in [0, \tau]$. This implies that the solution $t \to x(t)$ can be prolonged to $+\infty$, i.e., $I_x = \mathbb{R}^+$ and we have $\|x(t) - \xi(t)\| < 2R$, $\|x(t)\| < 3R$, for any $t \geq 0$. If $x(t) \in C(t)$ for any $t \in [0, \tau]$, then the result is true due to Assumption 1. Suppose instead there are $t \in [0, \tau]$ such that $x(t) \notin C(t)$. Since $\sigma > \sigma_a$, for a.e. $t \in [0, \tau]$ such that $x(t) \notin C(t)$, and $\|x(t) - \xi(t)\| \leq 2R$, we have $\|x(t)\| < 3R$ and on the basis of Lemma 2

$$\begin{aligned} \dot{\Delta}(t) &< -2R(M_F(3R) + L_\xi + h^2) + 2R(M_F(3R) + L_\xi) \\ &= -2Rh^2 < 0. \end{aligned}$$

We have $\|x(0) - \xi(0)\| < 2R$. Suppose for contradiction that there exists $\hat{t} \in [0, \tau]$ such that $\|x(\hat{t}) - \xi(\hat{t})\| = 2R$ and $\|x(t) - \xi(t)\| < 2R$ for any $t \in [0, \hat{t})$. Since $\xi(\hat{t}) \in B(0, R)$, then we have $\|x(\hat{t})\| > R$ and $x(\hat{t}) \notin C(\hat{t})$. Let $t_i = \sup\{0 < t < \hat{t} : \|x(t)\| \leq R\}$. We have $0 < t_i < \hat{t}$, $\|x(t_i)\| = R$ and so $\|x(t_i) - \xi(t_i)\| < 2R$, moreover $x(t) \notin C(t)$, $t \in [t_i, \hat{t}]$. This yields $\dot{\Delta}(t) < -2Rh^2 < 0$ for a.e. $t \in [t_i, \hat{t}]$. Then we have $\Delta(\hat{t}) < \Delta(t_i)$, i.e., $\|x(\hat{t}) - \xi(\hat{t})\| = 2R < \|x(t_i) - \xi(t_i)\| < 2R$, which is a contradiction. ■

Proposition 2: Suppose that the same assumptions as in Proposition 1 are satisfied. Suppose in addition that one of the following assumptions hold.

1) F is a continuous single-valued map and, for any $t \geq 0$, $x \to F(t, x)$ is Lipschitz;

2) for any $t \geq 0$ the multivalued map $x \to -F(t, x)$ is maximal monotone (cf. [41, Def. 1, p. 140]).

Then, for any $x_0 \in B(0, R)$ there exists a unique solution $x(t)$, $t \geq 0$, of (2) with initial condition $x(0) = x_0$.

Proof: We provide a sketch of the proof.

1) Since $x \to W(t, x)$ is convex for any $t \geq 0$, and $\sigma > 0$, $x \to \sigma\partial W(t, x)$ is a maximal monotone operator [41, Prop. 1, p. 159]. Then, arguing as in the proof of [46, Prop. 3, p. 3531], we can conclude that there exists a unique solution $x(t)$, $t \geq 0$, of (2) such that $x(0) = x_0$.

2) The map $x \to -F(t, x) + \sigma\partial_x W(t, x)$ is maximal monotone as the sum of two operators enjoying this property [47]. Arguing once more as in the proof of [46, Prop. 3, p. 3531], we obtain the stated uniqueness result. ■

B. Reaching phase

By using the distance $\Delta(\cdot)$ in (5), as a Lyapunov function for (2), we can also address the reaching phase of the moving set $C(\cdot)$, as stated in the next result.

Theorem 1: Suppose that Assumpions 1-3 are satisfied and that we have $\sigma > \sigma_a$, where σ_a is given in (6). Then, for any solution $x(t)$, $t \geq 0$, of NTN such that $x(0) \in B(0, R) \backslash C(0)$, there exists an instant

$$t_h \leq \frac{\|x(0) - \xi(0)\|^2}{4Rh^2} < \frac{R}{h^2} < +\infty \tag{9}$$

where h^2 is given in (8), such that we have $x(t_h) \in C(t_h)$, i.e., the solution reaches the moving set in finite time t_h.

Proof: Since $x(0) \notin C(0)$, due to continuity of $t \to x(t)$ and $t \to C(t)$ (Assumption 1), there exists $\delta > 0$ such that $x(t) \notin C(t)$, $t \in [0, \delta)$. Let $T_h = \sup\{\delta \geq 0 : x(t) \notin C(t), t \in [0, \delta)\}$, where $0 < T_h \leq +\infty$. By (7), for a.e. $t \in [0, T_h]$ we have $\dot{\Delta}(t) < -2Rh^2 < 0$ and so

$$\frac{1}{2}\|x(t) - \xi(t)\|^2 < \frac{1}{2}\|x(0) - \xi(0)\|^2 - 2Rh^2 t$$

for any $t \in [0, T_h]$. Since $C(t)$ is nonempty, compact and convex, and $\xi(t) \in \text{int}C(t)$, we have $\text{dist}(x(t), C(t)) < \|x(t) - \xi(t)\|$. By the previous inequality we have $T_h < \|x(0) - \xi(0)\|^2/(4Rh^2) < R/h^2$, where we noted that $\|x(0) - \xi(0)\| < 2R$. Then, there exists $t_h \leq T_h$ such that $x(t) \notin C(t)$ for any $t \in [0, t_h)$ and $x(t_h) \in C(t_h)$. ■

Theorem 1 guarantees that any solution of (2) starting in $B(0, R)\backslash C(0)$ reaches $C(\cdot)$ in finite time. However, as shown in the next example, it does not guarantee that the solution stays within the set thereafter. This may be a drawback in view of the application to the solution of TD constraint satisfaction problems (cf. Section V).

Example. Let $q = 2$ and $x = (x_1, x_2)' \in \mathbb{R}^2$, where the prime means transpose. Suppose $C(0)$ is a rectangle centered at the origin with vertexes $(6, 1)$, $(6, -1)$, $(-6, -1)$, $(-6, 1)$, and that $C(t)$ is obtained by a rigid rotation in clockwise sense of $C(0)$ by an angle $\theta(t)$ obeying the following law: we have $\dot{\theta}(t) = 0$, $t \in (0, 21)$ sec, whereas $\dot{\theta}(t) = \pi/3$ rad/sec, $t \in (21, 24)$ sec. Actually, $C(\cdot)$ is standing still in $(0, 21)$ sec and it undergoes a quick clockwise rotation of π radians in the interval $(21, 24)$ sec. Then, this behavior repeats with a period of 24 sec. Also assume that there is a biasing vector field $F(t, x) = (0.2, 0)'$. For brevity we omit to write the expressions of the inequality constraints defining $C(\cdot)$. It can be checked however that Assumptions 1-3 are satisfied by choosing $R = 7$ and $\xi(t) = 0$ for any $t \geq 0$. We have $L_\xi = 0$, $\psi = 1$, and then $\sigma_a = 2.8$.

Let $\sigma = 3$ and let $x(t)$, $t \geq 0$, be the unique solution of NTN starting at $x(0) = (-1, -4)' \in B(0, 7)\backslash C(0)$ at $t = 0$ (Proposition 2). It can be seen from MATLAB simulations that the trajectory first hits $C(\cdot)$ at an instant less than t_h in (9), in accordance with Theorem 1 (Fig. 1). However, thereafter it repeatedly enters and exits $C(\cdot)$. Eventually, the trajectory tends to a cycle which is in part contained in $C(\cdot)$ and in part located outside $C(\cdot)$. This behavior can be explained by noting that the penalty parameter σ_a in Theorem 1 does not account for the rotation of $C(\cdot)$. It is intuitively clear that the trajectories of (2) may not be able to follow fast rotations of $C(\cdot)$, and remain within the same set, unless σ is suitably chosen also in dependence of the rotation speed. This issue will be tackled in Theorem 2 of the next section.

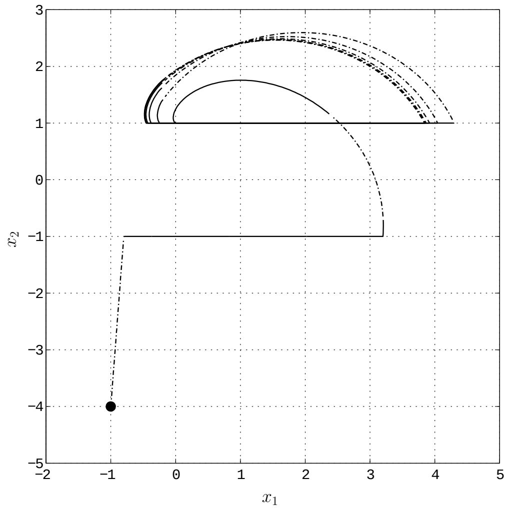

Fig. 1. Trajectory of NTN starting at $(-1,-4)$ when $\sigma = 3$. The dashed-dotted part represents points $x(t) \notin C(t)$, whereas the solid part corresponds to $x(t) \in C(t)$.

C. Tracking phase

In order to guarantee that once a solution of NTN reaches the moving set $C(\cdot)$, then it stays within the set thereafter, we use in what follows as a Lyapunov function for NTN the barrier function $W(t,x)$ in (1).

The next result is useful for evaluating the time derivative of $W(t,x)$ along the solutions of NTN.

Lemma 3: Suppose that Assumpions 1-3 are satisfied. Let $x(t)$, $t \geq 0$, be a solution of (2) such that $x(0) \in B(0,R)\backslash C(0)$. Suppose in addition that we have $x(t) \in B(0,R)$ for any $t \in [t_a,t_b] \subset \mathbb{R}^+$. Then, for a.e. $t \in [t_a,t_b]$ such that $x(t) \notin C(t)$ we have

$$\frac{d}{dt}W(t,x(t)) \leq \|z(t)\|(\|\gamma(t)\| - \sigma\|z(t)\|) + \sum_{j=1}^{p} L_j(R)$$

where $\gamma(t) \in F(t,x(t))$ and $z(t) \in \partial_x W(t,x(t))$.

Proof: Since $t \to x(t)$ is absolutely continuous on any compact interval in $\mathbb{R}^+$, and $(t,x) \to W(t,x)$ is Lipschitz with Lipschitz constant $\sum_{j=1}^{p} L_j(R)$, for a.e. $t \geq 0$ there exist both $\dot{x}(t)$ and

$$\begin{aligned}\frac{d}{dt}W(t,x(t)) &= \lim_{\tau\to 0}\frac{W(t+\tau,x(t+\tau)) - W(t,x(t))}{\tau}\\ &= \lim_{\tau\to 0}\Big[\frac{W(t+\tau,x(t+\tau)) - W(t,x(t+\tau))}{\tau}\\ &\quad + \frac{W(t,x(t+\tau)) - W(t,x(t))}{\tau}\Big].\end{aligned}$$

By assumption $x \to W(t,x)$ is convex for any $t \geq 0$. Then, by the Chain Rule, see point iii) of Section I-A2, we obtain

$$\lim_{\tau\to 0}\frac{W(t,x(t+\tau)) - W(t,x(t))}{\tau} = \langle u(t),\dot{x}(t)\rangle \quad \forall u(t) \in \partial_x W(t,x(t)).$$

As it can be easily verified, for a.e. $t \in [t_a,t_b]$, we have $|W(t+\tau,x(t+\tau)) - W(t,x(t+\tau))| = \big|\sum_{j=1}^{p}(\max\{0,-f_j(t+\tau,x(t+\tau))\} - \max\{0,-f_j(t,x(t+\tau))\})\big| \leq \tau\sum_{j=1}^{p} L_j(R)$, and then

$$\frac{d}{dt}W(t,x(t)) \leq \langle u(t),\dot{x}(t)\rangle + \sum_{j=1}^{p} L_j(R) \quad \forall u(t) \in \partial_x W(t,x(t)).$$

We have $\dot{x}(t) \in F(t,x(t)) - \sigma\partial_x W(t,x(t))$, hence there exists $\gamma(t) \in F(t,x(t))$ and $z(t) \in \partial_x W(t,x(t))$ such that $\dot{x}(t) = \gamma(t) - \sigma z(t)$. By taking $u(t) = z(t) \in \partial_x W(t,x(t))$ we obtain

$$\begin{aligned}\frac{d}{dt}W(t,x(t)) &\leq \langle z(t),\gamma(t)\rangle - \sigma\|z(t)\|^2 + \sum_{j=1}^{p} L_j(R)\\ &\leq \|z(t)\|(\|\gamma(t)\| - \sigma\|z(t)\|) + \sum_{j=1}^{p} L_j(R).\end{aligned}$$

■

Let

$$\rho_{\mathrm{M}} \doteq \sup_{t\geq 0}\sup_{y\in \mathrm{bd}C(t)}\|y - \xi(t)\|.$$

Note that under Assumptions 1, 2 we have $\rho_{\mathrm{M}} \leq 2R$.

Theorem 2: Suppose that Assumpions 1-3 are satisfied. Also let $\sigma > \max\{\sigma_a,\sigma_b\}$, where σ_a is given in (6) and

$$\sigma_b = \frac{\rho_{\mathrm{M}}}{\psi}\left(M_F(R) + \frac{\rho_{\mathrm{M}}}{\psi}\sum_{j=1}^{p} L_j(R)\right). \tag{10}$$

Then, any solution $x(t)$, $t \geq 0$, of (2) such that $x(0) \in B(0,R)\backslash C(0)$, is such that $x(t_h) \in C(t_h)$, and $x(t) \in C(t)$ for any $t \geq t_h$, where t_h is given in (9), i.e., the solution reaches the moving set in finite time and stays there thereafter.

Proof: Let $x(t)$, $t \geq 0$, be a solution of (2) such that $x(0) \in B(0,R)\backslash C(0)$. On the basis of Theorem 1 there exists t_h as in (9) such that $x(t_h) \in C(t_h)$ and so, by Assumption 1, $x(t_h) \in B(0,R)$. Suppose now for contradiction that there exists $\hat{t} > t_h$ such that $x(\hat{t}) \notin C(\hat{t})$. We can assume without loss of generality that we have $x(\hat{t}) \in B(0,R)$ and $x(t) \in B(0,R)$ for any $t \in [t_h,\hat{t}]$. Let $t_i = \sup\{t < \hat{t} : x(t) \in C(t)\}$. We have $t_h \leq t_i < \hat{t}$, $x(t_i) \in C(t_i)$ and $W(t_i,x(t_i)) = 0$.

Fix any $\sigma > \sigma_b$ and choose $0 < \epsilon < \sqrt{\sigma/\sigma_b} - 1$. Since $x(t_i) \in C(t_i)$, due to continuity of $t \to x(t)$ and the Lipschitz continuity of $t \to C(t)$, we can find an instant $\tilde{t} \in (t_i,\hat{t}]$ such that we have $\mathrm{dist}(x(t),C(t)) < \epsilon\rho_{\mathrm{M}}$ for any $t \in (t_i,\tilde{t})$ and then $\|x(t) - \xi(t)\| \leq \mathrm{dist}(x(t),C(t)) + \mathrm{dist}(C(t),\xi(t)) < \rho_{\mathrm{M}}(1+\epsilon)$ for any $t \in (t_i,\tilde{t})$. We also have $x(\tilde{t}) \notin C(\tilde{t})$ and then $W(\tilde{t},x(\tilde{t})) > 0$. We want to show that $dW(t,x(t))/dt < 0$ for a.e. $t \in [t_i,\tilde{t}]$ and so, by integrating in $[t_i,\tilde{t}]$, we obtain $0 = W(t_i,x(t_i)) > W(\tilde{t},x(\tilde{t})) > 0$, which is a contradiction. To show that $dW(t,x(t))/dt < 0$ for a.e. $t \in [t_i,\tilde{t}]$ we first note that, since $z(t) \in \partial_x W(t,x(t))$, and $x(t) \notin C(t)$ for any $t \in (t_i,\tilde{t})$, Lemma 1 yields $\|x(t) - \xi(t)\|\|z(t)\| \geq \langle x(t) - \xi(t),z(t)\rangle \geq \psi + W(t,x(t))$, and then

$$\|z(t)\| \geq \frac{\psi + W(t,x(t))}{\|x(t) - \xi(t)\|} > \frac{\psi}{\rho_{\mathrm{M}}(1+\epsilon)} > 0.$$

By Lemma 3 we obtain for a.e. $t \in [t_i,\tilde{t}]$

$$\frac{d}{dt}W(t,x(t)) \leq \|z(t)\|^2\big(-\sigma + \frac{\|\gamma(t)\|}{\|z(t)\|} + \frac{\sum_{j=1}^{p} L_j(R)}{\|z(t)\|^2}\big)$$

where $\gamma(t) \in F(t, x(t))$ and $z(t) \in \partial_x W(t, x(t))$ are such that $\dot{x}(t) = \gamma(t) - \sigma z(t)$. Therefore

$$\begin{aligned}\frac{d}{dt}W(t,x(t)) &< \|z(t)\|^2\big(-\sigma + \frac{\rho_{\mathrm{M}} M_F(R)(1+\epsilon)}{\psi} \\ &\quad + \frac{\rho_{\mathrm{M}}^2(1+\epsilon)^2\sum_{j=1}^p L_j(R)}{\psi^2}\big) \\ &< (1+\epsilon)^2\|z(t)\|^2\big(-\frac{\sigma}{(1+\epsilon)^2} \\ &\quad + \frac{\rho_{\mathrm{M}} M_F(R)}{\psi} + \frac{\rho_{\mathrm{M}}^2\sum_{j=1}^p L_j(R)}{\psi^2}\big)\end{aligned}$$

where we have considered that $x(t) \in B(0, R)$, $t \in (t_i, \tilde{t})$, and then $\|\gamma(t)\| \le M_F(R)$. We have $\sigma/(1+\epsilon)^2 > \sigma_b$, yielding $dW(t, x(t))/dt < 0$ for a.e. $t \in [t_i, \tilde{t}]$. ∎

D. Remarks

1) Theorem 2 gives an easy to compute estimate of the time t_h needed for a solution starting in $B(0,R)\backslash C(0)$, where R is given in Assumption 1, to reach the moving set $C(\cdot)$ and track the same set (cf. (9)). Thresholds σ_a, σ_b in Theorem 2 can be effectively evaluated on the basis of the parameters of NTN and the moving set, as discussed via specific examples in Section V. It is important to stress that: (a) the estimate of t_h in Theorem 2 is uniform with respect to initial conditions in $B(0,R)\backslash C(0)$; (b) the estimate of t_h is the same for all possible solutions starting at a given initial condition $x(0) \in B(0,R)\backslash C(0)$. As seen in Proposition 2, and in the examples in Section V, there are relevant classes of vector fields $F(t,x)$ for which the uniqueness of solutions of (2) with respect to initial conditions holds. However, the general treatment in the paper encompasses also situations where (2) has multiple solutions starting at $x(0)$ (an example of this kind can be obtained by adapting [13, Ex. 3, p. 1748] to the present TD setting). As pointed out before, the estimate t_h in Theorem 2 holds for any possible solution starting at $x(0)$, so that the application of the theorem is simple and direct also in the multiple solution case.

2) In recent years a new type of neural network, termed Zhang neural network (ZNN), has been developed for solving in real time certain classes of TD problems. ZNN has been successfully applied to solve TD Sylvester equations [48], to find the inverse of TD matrices [49] and, more recently, to solve certain TD linear matrix inequalities (LMIs) by means of an auxiliary equality matrix [11]. ZNN exploits a vector-valued error function, and the information on the time derivatives of the TD parameters or coefficients involved, for ensuring exponential convergence to the solution. It is noted that ZNN is described by a smooth system of differential equations and it is applicable to *smooth* TD problems where parameters and coefficients are differentiable functions of time.

The crucial aspect of the neural network NTN proposed in this paper is instead the *nonsmoothness* of the dynamics and the possibility to apply NTN to some classes of NS-TD problems. Nonsmoothness of NTN is important in at least three aspects. a) The result for NTN in Theorem 2 can be applied to convex TD constraint satisfaction problems where the constraints are in general nondifferentiable functions of the state variables or of time. b) Even for smooth constraints the NS dynamics of NTN plays a crucial role for ensuring convergence in *finite time* to the moving set, via the use of a NS-TD barrier function and an exact penalty function method, and also for ensuring an exact tracking of the moving set, through the sliding modes which are peculiar to the NS dynamics. c) We can handle by means of NTN the presence of very general NS and multi-valued vector fields F describing the unconstrained dynamics.

3) As pointed out by the Reviewers, it would be interesting to ask whether the proposed neural network (2) can be used also for TD constrained optimization problems. Namely, to design (2) in such a way that each solution reaches in finite time the set $M(t)$ of constrained minima of a given TD cost function $\phi(t,x)$ and stays within $M(t)$ thereafter. Here we first remark that TD optimization problems of this kind cannot be dealt with in their full generality by means of the results in this paper. The main reason is that, as it can be verified by means of simple examples, even if $\phi(t,x)$ is a smooth function, the set of constrained minima $t \to M(t)$ may be discontinuous with time, thus violating the requirement of Lipschitz continuity, as in iii) of Assumption 1, on which the present analysis is based. On the other hand, we believe that a challenging problem for future work is to see if there are subclasses of TD optimization problems than can be solved by the proposed network (2) or by suitably modifying its dynamics.

IV. Tracking Phase as a Moreau Sweeping Process

Sweeping processes were originally introduced by Moreau in the 70' as a framework for describing the evolution of a dynamical system under TD convex constraints [50]. Processes of this kind arise naturally in such areas as NS mechanics, convex optimization, econometrics, control, and many others [40], [51], [52]. Here we show that the tracking phase of the NTN dynamics can be interpreted as an instance of a sweeping process. We believe this aspect is of interest from a theoretic viewpoint since it shows that NS-TD neural networks as those proposed in this paper can effectively be used to model relevant TD nonlinear phenomena as sweeping processes.

As discussed in [40], in a sweeping process there are two main ingredients, i.e., a convex sweeping set $C(t)$ (e.g., a large ring) and a point (e.g., a smaller ball) that is located within $C(0)$ at $t = 0$ and is swept by the moving set $C(\cdot)$ for subsequent times $t \ge 0$. If $u(t) \in \mathbb{R}^q$ denotes the position of the point at time t, a simple model of a sweeping process is given by the differential inclusion

$$\dot{u} \in -N_{C(t)}(u), \quad u(0) \in C(0) \tag{11}$$

for a.e. $t \ge 0$, where $N_{C(t)}(u(t))$ denotes the normal cone to $C(t)$ at $u(t) \in C(t)$ (Section I-A3). From a geometric viewpoint, (11) means that the set $-N_{C(t)}(u(t))$ of admissible velocities at time t vanishes in the interior of $C(t)$ but points toward the interior of $C(t)$ on the boundary of $C(t)$. Then, the point does not move if it is in the interior of $C(\cdot)$ but it cannot exit $C(\cdot)$ and it is swept around by $C(\cdot)$ as soon as it touches the boundary of $C(\cdot)$. Equations of the type (11), where the normal cone prevents solutions from exiting the set, are also referred to in the literature as differential variational inequalities [41], [46], [53]. Sometimes also perturbed sweeping processes are considered where the right-hand side of (11) accounts for the presence of a vector field $F(t,u)$ modeling for instance external forces acting on the ball [54].

The analysis of sweeping processes aims at proving the existence of viable solutions, where a viable solution of (11) is by definition a function $u:\mathbb{R}^+\to\mathbb{R}^q$ such that: $u(0)\in C(0)$ and $u(t)\in C(t)$ for any $t\geq 0$; u is absolutely continuous on any compact interval of $\mathbb{R}^+$; we have $\dot{u}(t)\in -N_{C(t)}u(t)$ for a.e. $t\geq 0$. A typical result (see, e.g., Theorems 2, 3 in [40]) is that if $t\to C(t)$ is Lipschitz continuous in the Hausdorff metric then, for any $u_0\in C(0)$, there exists a unique viable solution of (11) with initial condition $u(0)=u_0$.

Consider now the simplified NTN model (2) where the forcing term $F(t,x)$ vanishes

$$\dot{x}\in -\sigma\partial_x W(t,x) \tag{12}$$

for a.e. $t\geq 0$. We can prove the following.

Theorem 3: Suppose that Assumptions 1-3 are satisfied and $\sigma>\max\{\sigma_a,\sigma_b\}$, where

$$\sigma_a=\frac{2R}{\psi};\;\; \sigma_b=\left(\frac{\rho_{\mathrm{M}}}{\psi}\right)^2\sum_{j=1}^{p}L_j(R).$$

Let $x_0\in B(0,R)\backslash C(0)$ and $x(t)$, $t\geq 0$, be the unique solution of (12) such that $x(0)=x_0$ (Proposition 2). Also let t_h as in (9) be such that $x(t_h)\in C(t_h)$ and $x(t)\in C(t)$ for any $t\geq t_h$ (Theorem 2). Then, $x(t)$ is the unique viable solution for $t\geq t_h$ of the TD DVI

$$\dot{x}\in -N_{C(t)}(x),\quad x(t_h)\in C(t_h).$$

Proof: The proof follows directly from Theorem 2 and the fact that if $x(t)\in \mathrm{bd}C(t)$, for some $t\geq t_h$, then $N_{C(t)}(t,x(t)) = \lim_{\sigma\to+\infty}\sigma\partial_x\sum_{j\in I^0(t,x(t))}\max\{0,-f_j(t,x(t))\} = \lim_{\sigma\to+\infty}\sigma\partial_x W(t,x(t))$, where $I^0(t,x(t)) = \{j\in\{1,2,\cdots,p\}: f_j(t,x(t))=0\}$. Observe that if $x(t)\in\mathrm{int}C(t)$ then $N_{C(t)}(t,x(t))=\{0\}$. ∎

Theorem 3 states that in the tracking phase any solution of (12) is the unique solution of a TD DVI modeling a sweeping process. In the proof of Theorem 3 we noted that $\sigma\partial_x W(t,x(t))\subset N_{C(t)}(x(t))$ for any $t\geq 0$. By the same proof, for $\sigma>\max\{\sigma_a,\sigma_b\}$, the term $\sigma\partial_x W(t,x(t))$ is seen to be a sufficiently large portion of $N_{C(t)}(x(t))$ forcing any solution of (2) to stay trapped within $C(\cdot)$ after the reaching time t_h. Then, the gradient of the NS-TD barrier function is a practical way to implement via neural networks the portion of the normal cone ensuring the existence of viable solutions that satisfy the TD convex constraints for $t\geq t_h$.

V. Applications and Examples

In this section we discuss the application of NTN to some toy examples of convex constraint satisfaction (feasibility) problems [55], [56]. Such problems can be stated as follows: suppose that $C_1,C_2,\ldots,C_p\subset\mathbb{R}^q$ are closed convex sets with nonempty intersection $C=C_1\cap C_2\ldots\cap C_p$; the (static) convex feasibility problem is simply to find a point $x\in C$.

As pointed out in the review paper [55], problems of this kind continue to receive great attention since they are very common and of extraordinary utility and broad applicability in many areas of mathematics, engineering and modern physical sciences. They arise for instance in solving partial differential equations, in statistics, complex analysis, medical imaging, in the set theoretic approach to image recovery, in signal processing and control theory, only to name a few of the main application fields. Many relevant convex feasibility problems are high-dimensional and are formulated in a time-varying environment. Moreover, providing their solution in a very quick time is crucial, as in fast signal encryption for the security of multimedia data [38], positioning in wireless sensor networks [57], image reconstruction from projections in computerized tomography [58], sensor placement for dense sampling in autonomous robotics [59], secure communications using robust masked beamforming schemes [60]).

The standard approach for solving convex feasibility problems is algorithmic and is based on projections methods [55]. An alternate approach, which we propose in this paper, is to exploit the real-time processing capabilities of dynamic neural networks. In the examples that follow we show that by using NTN we are able to address even TD versions of convex feasibility problems where we allow the sets $C_1(t),C_2(t),\ldots,C_p(t)$, and their intersection $C(t)$, to be time-varying. In particular we show that NTN is able to find in finite time t_h an exact feasible solution $x(t_h)\in C(t_h)$ and also give an exact feasible solution $x(t)\in C(t)$ for all subsequent times $t>t_h$. In Section V-A we first consider the application of NTN to problems where the convex sets are defined via affine inequality constraints, and there is an affine vector field, whereas in Section V-B we illustrate the application to sets defined by more general convex inequality constraints and vector fields.

A. Affine constraints and vector field

Here we consider the relevant case of affine inequality constraints $f(t,x)=(f_1(t,x),f_2(t,x),\ldots,f_p(t,x))'=B(t)x-e(t):\mathbb{R}^+\times\mathbb{R}^q\to\mathbb{R}^p$, where $B(\cdot)$ is a TD $p\times q$ matrix which is Lipschitz in $\mathbb{R}^+$ and $e(\cdot)$ is TD p-vector which is Lipschitz in $\mathbb{R}^+$. We also suppose that the vector field is affine, i.e., we have $F(t,x)=G(t)x+h(t)$, where $G(\cdot)$ is a TD $q\times q$ matrix which is Lipschitz for $t\geq 0$ and $h(\cdot)$ is TD q-vector which is Lipschitz for $t\geq 0$. Note that, under Assumption 1, $P(t)=C(t)=\{x\in\mathbb{R}^q: B(t)x-e(t)\geq 0\}$ is a polyhedron with nonempty interior. Also note that we have $P(\cdot)=H_1(t)\cap H_2(t)\ldots\cap H_p(t)$ where, for $j=1,2,\ldots,p$, $H_j(t)=\{x\in\mathbb{R}^q:\langle x,B_j(t)\rangle\geq e_j(t)\}$ is a half-plane in $\mathbb{R}^q$ with normal $B_j(t)$ corresponding to the j-th row of $B(t)$. The examples that follow can then be interpreted as TD convex feasibility problems.

For affine inequality constraints and vector field the NTN equations become

$$\dot{x}\in G(t)x+h(t)-B'(t)D(B(t)x-e(t)) \tag{13}$$

where we have let $D(y) = (\hat{d}(y_1), \hat{d}(y_2), \ldots, \hat{d}(y_n))'$. Note that, due to Proposition 2, (13) enjoys the uniqueness of solutions with respect to initial conditions.

In Section V-A1 we will provide three examples in the case where matrices B, G are constant and then, in Section V-A2, we discuss an example in the more general case where we have TD matrices $B(\cdot), G(\cdot)$. In the MATLAB simulations of these networks we have approximated the graph of the discontinuous function $\hat{d}$ in (3) with a piecewise linear Lipschitz function with high slope σ/ϵ given by $\sigma\hat{s}(\rho)$, where $\hat{s}(\rho) = 0$ for $\rho > 0$, $\hat{s}(\rho) = \rho/\epsilon$ for $\sigma \in [-\epsilon, 0]$, $\hat{s}(\rho) = -1$ for $\rho < -\epsilon$, and ϵ is a small positive number actually set to 0.01. This is justified since solutions of actual systems with high-gain nonlinearities are approximations, in the sense of the uniform norm, of solutions in the sense of Filippov as those considered in the paper. The reader is referred to [41, Th. 3, p. 98] for results on the approximation of Filippov solutions in high-gain systems, and to [61, Sect. II-C] for approximation in graph of high-gain nonlinearities used for modeling neural networks.

1) Constant interconnection matrices: Suppose the inequality constraints are given by $f(t, x) = Bx - e(t)$, where $B \in \mathbb{R}^{p\times q}$ is a constant matrix, and the vector field $F(t, x) = Gx + h(t)$, where $G \in \mathbb{R}^{q\times q}$. The NTN equations become

$$\dot{x} \in Gx + h(t) - B'D(Bx - e(t)). \tag{14}$$

In this case NTN admits a simple electronic implementation which is analogous to that in [2, Fig. 6]. NTN has two sets of neurons (state-variable and inequality-constraint neurons), which are interconnected via matrices B, G, and two biasing inputs $e(t), h(t)$. The entries of B, G can be implemented by means of constant conductances. Each state-variable neuron can be implemented by means of a linear summing operational amplifier, whereas each inequality-constraint neuron can be implemented by an operational amplifier with a diode in its feedback path [13]. The TD biases $e(t), h(t)$ can be realized by TD voltage sources as input signals to the network.

Since B is a constant matrix, $P(t) = \{x \in \mathbb{R}^q : Bx - e(t) \geq 0\}$ is a moving polyhedron such that each face moves parallel to itself on the basis of the changes in time of the biasing terms $e_i(t)$, $i = 1, 2, \ldots, p$. As shown in the next examples, by suitably choosing the TD biases we can implement various types of motion of $P(\cdot)$.

Example 1. Suppose to choose $e(t) = \bar{e} - B\zeta(t)$, where $\bar{e} \in \mathbb{R}^p$ is a constant vector and $\zeta(t)$ is a q-vector which is Lipschitz for $t \geq 0$. Consider the polyhedron $\bar{P} = \{x \in \mathbb{R}^q : Bx - \bar{e} \geq 0\}$. We have $P(t) = \bar{P} + \zeta(t)$, i.e., $P(t)$ is a polyhedron that rigidly translates driven by point $\zeta(t)$.

Suppose in particular that $q = 2, G = 0, h(t) = 0$ (no vector field), and there are $p = 4$ inequality constraints defined by

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & -1 \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

and $e(t) = (-1, 0, -1, -\sqrt{2})' - B(3\cos(t), 3\sin(t))' = \bar{e} - B\zeta(t)$. Note that $\bar{P} = \{x \in \mathbb{R}^2 : Bx - \bar{e} \geq 0\}$ is a trapezoidal set with vertexes $(2, 0)$, $(-1, 0)$, $(-1, 1)$, $(1, 1)$

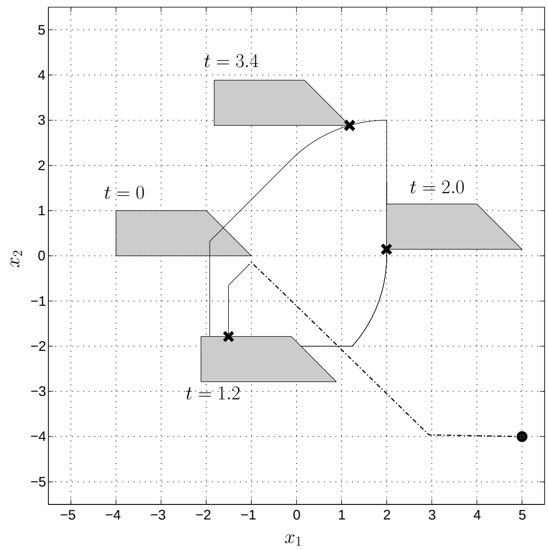

Fig. 2. Trajectory starting at $(5, -4)$ when $\sigma = 125$ for the NTN of Example 1. The trajectory hits the moving set $P(\cdot)$ in finite time and stays within the set thereafter. The dashed-dotted part represents points $x(t) \notin P(t)$, whereas the solid part corresponds to $x(t) \in P(t)$. The figure displays snapshots of $P(\cdot)$ at four different instants $t = 0$, $t = 1.2$, $t = 2$, $t = 3.4$ sec and points $x(t = 1.2) \in P(t = 1.2)$, $x(t = 2) \in P(t = 2)$, $x(t = 3.4) \in P(t = 3.4)$ denoted by crosses.

and $P(t) = \bar{P} + \zeta(t)$, i.e., $\bar{P}$ rigidly translates driven by $\zeta(t) = (3\cos(t), 3\sin(t))'$, which displays a circular motion around the origin. The biasing terms $e_i(t)$, $i = 1, 2, 3, 4$, can be implemented by batteries and sinusoidal voltage sources.

Let us choose $\xi(t) = \zeta(t) + (1/2, 1/2)'$. Since $t \to \xi(t)$ is Lipschitz with Lipschitz constant $L_\xi = 3$, also $t \to P(t)$ is Lipschitz continuous with the same Lipschitz constant $L_C = L_\xi = 3$ [40, Ex. 2, p. 9]. We can see that Assumption 1 is satisfied with $R = 7$ and $L_\xi = 3$ and also Assumption 2 is satisfied with $\psi = 1/2$. We obtain in particular $\sigma_a = 84$. Note that we can choose the Lipschitz constants L_j independent of R, namely, $L_j = 3$, $j = 1, 2, 3, 4$. By considering that $\rho_{\mathrm{M}} = \sqrt{10}/2$ we obtain $\sigma_b = 120$. Figure 2 shows the trajectory of NTN (14), starting at $(5, -4)$ at $t = 0$, when $\sigma = 125$. It is seen that the trajectory reaches $P(\cdot)$ in finite time and stays within $P(\cdot)$ thereafter, in accordance with Theorem 2. In the long-run behavior the trajectory tends to a cycle contained in $P(\cdot)$. An analogous behavior has been observed for other trajectories starting at different initial conditions.

We remark that, as it often happens in penalty function methods, the penalty parameter estimate is quite conservative since it is based on analyzing worst case situations. By reconsidering the proof of Theorem 2 it can be seen that the estimate can be remarkably improved when more information is available on the structure of the inequality constraints. For instance, in Theorem 2 we can substitute σ_b in (10) with

$$\sigma'_b = \frac{\rho_{\mathrm{M}}}{\psi}\left(M_F(R) + \frac{\rho_{\mathrm{M}}}{\psi}p'L(R)\right)$$

where p' is the maximal number of simultaneously active constraints in a small neighborhood of the moving set and $L(R) = \max_{j=1,2,\ldots,p} L_j(R)$. In the current example we have $p' = 2 < p = 4$ and $L = L(R) = 3$, yielding $\sigma'_b = 60$.

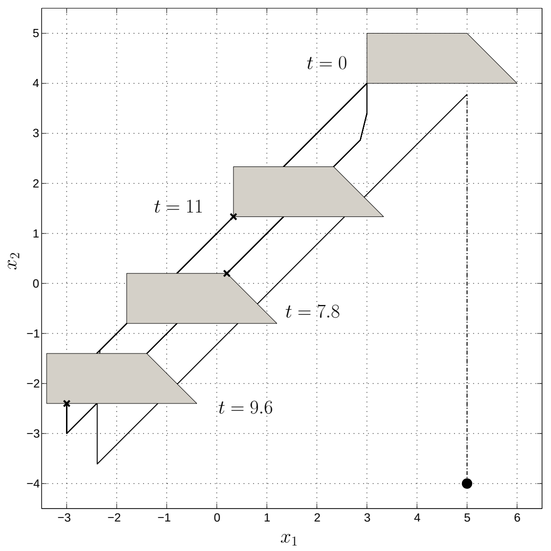

Fig. 3. Snapshots of the moving set and trajectory starting at $(5,-4)$ for the NTN of Example 2.

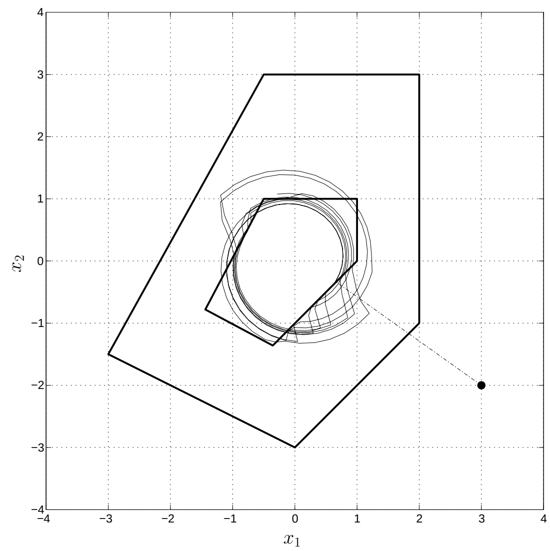

Fig. 4. Limiting sets and trajectory starting at $(3,-2)$ for the NTN of Example 3.

From MATLAB simulations it appears that the tracking and reaching phase still works for σ as low as 15.

Example 2. Consider the same problem as in Example 1 but suppose now that $\zeta(\cdot)$ displays an alternate motion with constant velocity (in modulus), along the segment with vertexes $(4,4)$, $(-4,-4)$, with a period of 6 sec. If $R=7$ and $\xi(t)=\zeta(t)+(1/2,1/2)'$, we obtain $\sigma_a=93.3$ and $\sigma_b'=66.6$. Figure 3 displays the trajectory of NTN (14) starting at $(5,-4)$ at $t=0$ for $\sigma=95$. Once more, in agreement with Theorem 2, the trajectory hits in finite time the moving set and stays in the set for subsequent times. Note that in this case the constraints f_j are nondifferentiable functions of time at the instants where there is an inversion of velocity, i.e., when $\zeta(\cdot)$ is on vertexes $(4,4)$, $(-4,-4)$. The alternate motion of $\zeta(\cdot)$ can be implemented using batteries and sawtooth voltage sources for the biasing terms $e_i(t)$, $i=1,2,3,4$.

Example 3. Suppose to choose $P(t)=\{x\in\mathbb{R}^q: Bx-e(t)\geq 0\}$ and assume in addition that $e_i(t)<0$, $i=1,2,\ldots,p$, for any $t\geq 0$. As it was noted before, $P(t)$ is a polyhedron that changes its shape in time since its faces independently move parallel to themselves. The assumption $e_i(t)<0$ for $i=1,2,\ldots,p$ and $t\geq 0$ ensures that $P(t)$ has nonempty interior and $0\in\operatorname{int}P(t)$ for any $t\geq 0$.

In particular, let $q=2$, $p=5$,

$$B=\begin{pmatrix} 0 & -1\\ -1 & 0\\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\\ \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5}\\ \frac{2\sqrt{5}}{5} & -\frac{\sqrt{5}}{5}\end{pmatrix}$$

and $e(t)=(-2+\cos(t),-1.5+0.5\sin(2t),-\sqrt{2}+\frac{\sqrt{2}}{2}\cos(3t),-\frac{9}{10}\sqrt{5}+\frac{3}{10}\sqrt{5}\sin(2t),-\frac{3}{5}\sqrt{5}-\frac{\sqrt{5}}{5}\cos(3t))'$. We also let $F(t,x)=Gx$ where

$$G=\begin{pmatrix} 0.1 & -0.9\\ 0.9 & 0.1\end{pmatrix}.$$

The set $P(\cdot)$ shrinks and expands between the two limiting sets shown in Fig. 4. If we let $R=4$ and $\xi(t)=0$, $t\geq 0$, hence $L_\xi=0$, we obtain $\sigma_a=124$, $\sigma_b'=129$. Figure 4 displays the trajectory of NTN (14) starting at $(3,-2)$ at $t=0$ when $\sigma=130$. It can be checked that the reaching and tracking behavior agrees with that predicted by Theorem 2.

2) TD interconnection matrices: We discuss an example showing that by using a TD matrix $B(t)$ for the affine inequality constraints we can implement more complex motions of $P(t)=\{x\in\mathbb{R}^q: B(t)x-e(t)\geq 0\}$ such as a roto-translational motion.

Example 4. Consider a parallelotope $\mathcal{P}(\mathcal{T},c)=\{x\in\mathbb{R}^q: x=\mathcal{T}\alpha+c; \|\alpha\|_\infty\leq 1\}$, where $\mathcal{T}\in\mathbb{R}^{q\times q}$ is the nonsingular matrix of the edges and $c\in\mathbb{R}^q$ is the center [62]. If the center is time-varying, then $\mathcal{P}(\mathcal{T},c(t))$ rigidly translates driven by point $c(t)$. If in addition $\hat{\mathcal{T}}(t)=\mathcal{R}(t)\mathcal{T}$, where $\mathcal{R}(t)\in\mathbb{R}^{q\times q}$ is a rotation matrix, then $\mathcal{P}(\hat{\mathcal{T}}(t),c(t))$ is a roto-translating parallelotope. By using the representation in this paper we have $\mathcal{P}(\hat{\mathcal{T}}(t),c(t))=\{x\in\mathbb{R}^q: B(t)x-e(t)\geq 0\}$ where

$$B(t)=\begin{pmatrix}\mathcal{T}^{-1}\mathcal{R}'(t)\\ -\mathcal{T}^{-1}\mathcal{R}'(t)\end{pmatrix}, e(t)=\begin{pmatrix}\mathcal{T}^{-1}\mathcal{R}'(t)c(t)-\hat{1}\\ -\mathcal{T}^{-1}\mathcal{R}'(t)c(t)-\hat{1}\end{pmatrix}$$

where $\hat{1}\in\mathbb{R}^q$ is a vector of 1's.

In particular, suppose $q=2$ and $G=0, h(t)=0$, and let

$$\mathcal{T}=\begin{pmatrix} 2 & 0\\ 0 & 1\end{pmatrix}, \mathcal{R}(t)=\begin{pmatrix}\cos(2t) & \sin(2t)\\ -\sin(2t) & \cos(2t)\end{pmatrix}$$

and $c(t)=(3\cos(t/2),0)'$. Then, we obtain

$$B(t)=\frac{1}{2}\begin{pmatrix}\cos(2t) & -2\sin(2t)\\ \sin(2t) & 2\cos(2t)\\ -\cos(2t) & 2\sin(2t)\\ -\sin(2t) & -2\cos(2t)\end{pmatrix}$$

and $e(t)=(1.5\cos(2t)\cos(t/2)-1, 3\sin(2t)\sin(t/2)-1, -1.5\cos(2t)\cos(t/2)-1, -3\sin(2t)\sin(t/2)-1)'$.

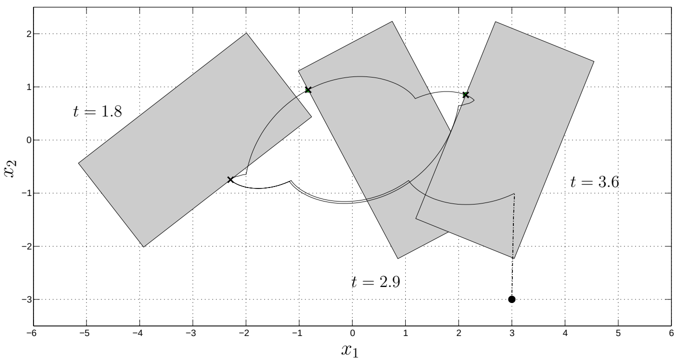

Fig. 5. Snapshots of the moving set and trajectory starting at $(3, -3)$ for the NTN of Example 4.

At $t = 0$ we have a rectangle with vertexes $(2,1),(2,-1),(-2,-1),(-2,1)$. The rectangle rotates around the center with angular speed 2 rad/sec, and the center moves along the x_1-axis with a law $x_1(t) = 3\cos(t/2)$. If $R = 7$ and $\xi(t) = c(t)$, we obtain $\sigma_a = 21$ and $\sigma_b' = 59.7$. Figure 5 shows the trajectory of (13) starting at $x(0) = (3,-3)'$ when $\sigma = 80$. Once more we can observe a reaching and tracking phase of the roto-translating rectangle in agreement with that predicted by Theorem 2. In this application the NTN (13) admits an electronic implementation similar to that discussed in Section V-A1. However, since matrix $B(\cdot)$ is TD, we need to use TD conductances, which can be realized for example using electronic devices as operational transconductance amplifiers.

B. More General Constraints and Vector Field

We simulated using MATLAB other examples where NTN is used for solving convex feasibility problems involving NS, non-affine, constraints f_j and vector fields F. In what follows we present one such example involving an inequality constraint depending quadratically in the state variables.

Example 5. Let $q = 2$. Suppose $F(t,x) = (0, 1/2)'$ and there are $p = 2$ inequality constraints

$$\begin{aligned} f_1(t,x) &= -\frac{4}{27}x_1^2 - \frac{1}{3}x_2^2 + \frac{4}{3} \geq 0 \\ f_2(t,x) &= -x_1\sin(t/2) + x_2\cos(t/2) \geq 0. \end{aligned}$$

The first one is an ellipse in $\mathbb{R}^2$ whose center is located at the origin. The second constraint is a rotating half-plane delimited by a rotating straight line pivoted at the origin.

The NTN equations are given as

$$\begin{aligned} \dot{x}_1 &\in -\hat{d}(f_1(t,x))\left(-\frac{8}{27}x_1\right) - \hat{d}(f_2(t,x))(-\sin(t/2)) \\ \dot{x}_2 &\in -\hat{d}(f_1(t,x))\left(-\frac{2}{3}x_2\right) - \hat{d}(f_2(t,x))(\cos(t/2)). \quad (15) \end{aligned}$$

If we choose $\xi(t) = (-\sin(t/2), \cos(t/2))'$ and $R = 5$ we obtain $\sigma_a = 10$ and $\sigma_b = 75$. Figure 6 shows the trajectory of NTN (15) starting at $(x_1(0), x_2(0)) = (3,-3)$ when $\sigma = 80$. Also in this case the trajectory hits the set $C(\cdot)$ in finite and remains trapped within the set thereafter. The implementation of NTN now requires 8 four-quadrant multipliers.

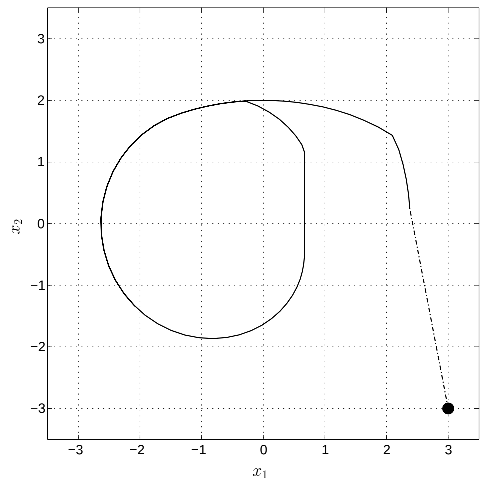

Fig. 6. Trajectory starting at $(3, -3)$ for the NTN of Example 5.

VI. Conclusion

The present paper is, to the authors' knowledge, the first investigation on NS neural network architectures for solving NS-TD problems. The proposed NTN has been designed via a NS-TD exact penalty function method. The theoretic investigation has shown that for a suitably chosen penalty parameter NTN displays a dynamics with two main phases, i.e., the reaching and tracking phase of a moving convex set. Simulations have shown that NTN is well suited for solving some classes of convex NS-TD constraint satisfaction problems, where NTN is able to find an exact feasible solution in finite time and provide an exact feasible solution for subsequent times. Other potential applications of the dynamics of NTN, which will be explored in future works, are in the field of tracking of moving objects or modeling of nonlinear dynamical systems displaying sweeping processes.

Acknowledgement

The authors thank to the Associate Editor and the anonymous Reviewers for the insightful and constructive comments, which helped to improve the presentation of the paper.

References

[1] D. W. Tank and J. J. Hopfield, "Simple 'neural' optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit," *IEEE Trans. Circuits Syst.*, vol. 33, pp. 533–541, May 1986.

[2] M. P. Kennedy and L. O. Chua, "Neural networks for nonlinear programming," *IEEE Trans. Circuits Syst. I*, vol. 35, pp. 554–562, May 1988.

[3] S. Zang and A. G. Constantinides, "Lagrange programming neural networks," *IEEE Trans. Circuits Syst. II*, vol. 39, pp. 441–452, Jul. 1992.

[4] A. Cichocki and R. Unbehauen, *Neural Networks for Optimization and Signal Processing*. Chichester, U.K.: Wiley, 1993.

[5] M. Forti and A. Tesi, "New conditions for global stability of neural networks with application to linear and quadratic programming problems," *IEEE Trans. Circuits Syst. I*, vol. 42, pp. 354–366, Jul. 1995.

[6] Y. Xia and J. Wang, "A general methodology for designing globally convergent optimization neural networks," *IEEE Trans. Neural Networks*, vol. 9, pp. 1331–1343, Nov. 1998.

[7] X. Hu and J. Wang, "Solving pseudomonotone variational inequalities and pseudoconvex optimization problems using the projection neural network," *IEEE Trans. Neural Networks*, vol. 17, pp. 1487–1499, Nov. 2006.

[8] Z. Zeng and J. Wang, "Improved conditions for global exponential stability of recurrent neural networks with time-varying delays," *IEEE Trans. Neural Networks*, vol. 17, pp. 623–635, May 2006.

[9] L. Cheng, Z.-G. Hou, and M. Tan, "A delayed projection neural network for solving linear variational inequalities," *IEEE Trans. Neural Networks*, vol. 20, pp. 915–925, Jun. 2009.

[10] Y. Xia, G. Feng, and J. Wang, "A novel recurrent neural network for solving nonlinear optimization problems with inequality constraints," *IEEE Trans. Neural Networks*, vol. 19, pp. 1340–1353, Aug. 2008.

[11] D. Guo and Y. Zhang, "Zhang neural network for online solution of time-varying linear matrix inequality aided with an equality conversion," *IEEE Trans. Neural Networks and Learning Systems*, vol. 25, pp. 370–382, Feb. 2014.

[12] Y. Xia, "A compact cooperative recurrent neural network for computing general constrained L_1 norm estimators," *IEEE Trans. Signal Processing*, vol. 57, pp. 3693–3697, Sep. 2009.

[13] M. Forti, P. Nistri, and M. Quincampoix, "Generalized neural network for nonsmooth nonlinear programming problems," *IEEE Trans. Circuits Syst. I*, vol. 51, pp. 1741–1754, Sep. 2004.

[14] Q. Liu and J. Wang, "A one-layer projection neural network for nonsmooth optimization subject to linear equalities and bound constraints," *IEEE Trans. Neural Networks and Learning Systems*, vol. 24, pp. 812–824, May 2013.

[15] G. Li, Z. Yan, and J. Wang, "A one-layer recurrent neural network for constrained nonsmooth invex optimization," *Neural Networks*, vol. 50, pp. 79–89, Feb. 2014.

[16] M. Forti, P. Nistri, and M. Quincampoix, "Convergence of neural networks for programming problems via a nonsmooth Lojasiewicz inequality," *IEEE Trans. Neural Networks*, vol. 17, pp. 1471–1486, Nov. 2006.

[17] W. Bian and X. Chen, "Neural network for nonsmooth, nonconvex constrained minimization via smooth approximation," *IEEE Trans. Neural Networks and Learning Systems*, vol. 25, pp. 545–556, Mar. 2014.

[18] D. Wang and L. Huang, "Periodicity and global exponential stability of generalized Cohen-Grossberg neural networks with discontinuous activations and mixed delays," *Neural Networks*, vol. 51, pp. 80–95, Mar. 2014.

[19] W. Bian and X. Xue, "Neural network for solving constrained convex optimization problems with global attractivity," *IEEE Trans. Circuits Syst. I*, vol. 60, pp. 710–723, Mar. 2013.

[20] Q. Liu and J. Wang, "A one-layer recurrent neural network for constrained nonsmooth optimization," *IEEE Trans. Syst., Man, and Cyber., Part B*, vol. 41, pp. 1323–1333, Oct. 2011.

[21] ——, "A one-layer recurrent neural network with a discontinuous hard-limiting activation function for quadratic programming," *IEEE Trans. Neural Networks*, vol. 19, pp. 558–570, Apr. 2008.

[22] X. Hu and J. Wang, "Design of general projection neural networks for solving monotone linear variational inequalities and linear and quadratic optimization problems," *IEEE Trans. Syst., Man, and Cyber., Part B*, vol. 37, pp. 1414–1421, Oct. 2007.

[23] K. Yuan and J. Cao, "An analysis of global asymptotic stability of delayed Cohen-Grossberg neural networks via nonsmooth analysis," *IEEE Trans. Circuits Syst. I*, vol. 52, pp. 1854–1861, Sep. 2005.

[24] W. Bian and X. Chen, "Smoothing neural network for constrained non-lipschitz optimization with applications," *IEEE Trans. Neural Networks and Learning Systems*, vol. 23, pp. 399–411, Mar. 2012.

[25] L. Cheng, Z.-G. Hou, Y. Lin, M. Tan, W. Zhang, and F.-X. Wu, "Recurrent neural network for non-smooth convex optimization problems with application to the identification of genetic regulatory networks," *IEEE Trans. Neural Networks*, vol. 22, pp. 714–726, May 2011.

[26] S. Qin, W. Bian, and X. Xue, "A new one-layer recurrent neural network for nonsmooth pseudoconvex optimization," *Neurocomputing*, vol. 120, pp. 655–662, Nov. 2013.

[27] R. Klein, "Project scheduling with time-varying resource constraints," *Int. J. Production Research*, vol. 38, no. 16, pp. 3937–3952, 2000.

[28] B.-S. Kang, G.-J. Park, and J. Arora, "A review of optimization of structures subjected to transient loads," *Structural and Multidisciplinary Optimization*, vol. 31, no. 2, pp. 81–95, Feb. 2006.

[29] H.-K. Chen, C.-F. Hsueh, and M.-S. Chang, "The real-time time-dependent vehicle routing problem," *Transp. Research Part E: Logistics and Transp. Review*, vol. 42, no. 5, pp. 383–408, Sep. 2006.

[30] N. Kuschel and R. Rackwitz, "Optimal design under time-variant reliability constraints," *Structural Safety*, vol. 22, no. 2, pp. 113–127, Jun. 2000.

[31] H. Doran and A. Rambaldi, "Applying linear time-varying constraints to econometric models: With an application to demand systems," *Journal of Econometrics*, vol. 79, no. 1, pp. 83–95, Jul. 1997.

[32] F. Belkhouche, "Reactive path planning in a dynamic environment," *IEEE Trans. Robotics*, vol. 25, pp. 902–911, Aug. 2009.

[33] B. C. Dean, "Algorithms for minimum-cost paths in time-dependent networks with waiting policies," *Networks*, vol. 44, no. 1, pp. 41–46, Aug. 2004.

[34] E. Bertolazzi, F. Biral, and M. Da Lio, "Real-time motion planning for multibody systems: Real life application examples," *Multibody System Dynamics*, vol. 17, no. 2-3, pp. 119–139, Apr. 2007.

[35] K. Tee, B. Ren, and S. Ge, "Control of nonlinear systems with time-varying output constraints," *Automatica*, vol. 47, no. 11, pp. 2511–2516, Nov. 2011.

[36] M. Tagmouti, M. Gendreau, and J.-Y. Potvin, "Arc routing problems with time-dependent service costs," *European J. Operational Research*, vol. 181, no. 1, pp. 30–39, Aug. 2007.

[37] H. Doran, "Estimation under exact linear time-varying constraints, with an application to population projections," *Journal of Forecasting*, vol. 15, no. 7, pp. 527–541, Dec. 1996.

[38] X. Yi, C. Tan, C. Siew, and M. Syed, "Fast encryption for multimedia," *IEEE Trans. Consumer Electronics*, vol. 47, pp. 101–107, Jan. 2001.

[39] Z. Wan and M. Kothare, "Efficient robust constrained model predictive control with a time varying terminal constraint set," *Systems and Control Letters*, vol. 48, no. 5, pp. 375–383, Apr. 2003.

[40] M. Kunze and M. D. Marques, *An Introduction to Moreau's Sweeping Process*, ser. Lecture Notes in Physics, B. Brogliato, Ed. Berlin, Heidelberg: Springer-Verlag, 2000.

[41] J. P. Aubin and A. Cellina, *Differential Inclusions*. Berlin: Springer-Verlag, 1984.

[42] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms I: Part 1: Fundamentals*. Berlin, Heidelberg: Springer-Verlag, 1996.

[43] M. Forti, M. Grazzini, P. Nistri, and L. Pancioni, "Generalized Lyapunov approach for convergence of neural networks with discontinuous or non-Lipschitz activations," *Physica D: Nonlinear Phenomena*, vol. 214, no. 1, pp. 88–99, Feb. 2006.

[44] J. P. Aubin and H. Frankowska, *Set-Valued Analysis*. Boston: Birkauser, 1990.

[45] M. D. Wills, "Hausdorff distance and convex sets," *Journal of Convex Analysis*, vol. 14, no. 1, pp. 109–117, 2007.

[46] M. Di Marco, M. Forti, M. Grazzini, P. Nistri, and L. Pancioni, "Lyapunov method and convergence of the full-range model of CNNs," *IEEE Trans. Circuits Syst. I*, vol. 55, pp. 3528–3541, Dec. 2008.

[47] R. Rockafellar, "On the maximality of sums of nonlinear monotone operators," *Trans. Amer. Math. Soc.*, vol. 149, no. 1, pp. 75–88, May 1970.

[48] Y. Zhang, D. Jiang, and J. Wang, "A recurrent neural network for solving Sylvester equation with time-varying coefficients," *IEEE Trans. Neural Networks*, vol. 13, pp. 1053–1063, Sep. 2002.

[49] Y. Zhang and S. S. Ge, "Design and analysis of a general recurrent neural network model for time-varying matrix inversion," *IEEE Trans. Neural Networks*, vol. 16, pp. 1477–1490, Nov. 2005.

[50] J. Moreau, "Evolution problem associated with a moving convex set in a Hilbert space," *Journal of Differential Equations*, vol. 26, no. 3, pp. 347–374, Dec. 1977.

[51] ——, "Numerical aspects of the sweeping process," *Computer Methods in Applied Mechanics and Engineering*, vol. 177, no. 3-4, pp. 329–349, Jul. 1999.

[52] G. Colombo, R. Henrion, N. Hoang, and B. Mordukhovich, "Optimal control of the sweeping process," *Dynamics of Continuous, Discrete and Impulsive Systems Series B: Applications and Algorithms*, vol. 19, no. 1-2, pp. 117–159, 2012.

[53] J. Pang and D. Stewart, "Differential variational inequalities," *Math. Programming, Series A*, vol. 113, no. 2, pp. 345–424, Jun. 2008.

[54] J. Edmond and L. Thibault, "BV solutions of nonconvex sweeping process differential inclusion with perturbation," *Journal of Differential Equations*, vol. 226, no. 1, pp. 135–179, Jul. 2006.

[55] H. Bauschke and J. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM Review*, vol. 38, no. 3, pp. 367–426, Sep. 1996.

[56] P. Combettes and P. Bondon, "Hard-constrained inconsistent signal feasibility problems," *IEEE Trans. Signal Processing*, vol. 47, pp. 2460–2468, Sep. 1999.

[57] M. R. Gholami, H. Wymeersch, E. G. Ström, and M. Rydström, "Wireless network positioning as a convex feasibility problem," *EURASIP J. Wireless Comm. Netw.*, vol. 2011, no. 1, pp. 1–15, Nov. 2011.

[58] J. Montagnat and H. Delingette, "4D deformable models with temporal constraints: Application to 4D cardiac image segmentation," *Medical Image Analysis*, vol. 9, no. 1, pp. 87–100, Feb. 2005.

[59] V. Bhatawadekar, D. Fehr, V. Morellas, and N. Papanikolopoulos, “A dynamic sensor placement algorithm for dense sampling,” in *Proc. 14th Int. Conf. Information Fusion (FUSION)*, Chicago, IL, March 2011, pp. 1–7.

[60] Y. Tang, W. Li, D. Ma, X. Zhang, and J. Wei, “Worst-case robust masked beamforming for secure broadcasting,” in *Proc. 2013 IEEE Wireless Communications and Networking Conference (WCNC)*, Shangai, China, Apr. 2013, pp. 4186–4191.

[61] G. De Sandre, M. Forti, P. Nistri, and A. Premoli, “Dynamical analysis of full-range cellular neural networks by exploiting differential variational inequalities,” *IEEE Trans. Circuits Syst. I*, vol. 54, pp. 1736–1749, Aug. 2007.

[62] L. Chisci, A. Garulli, and G. Zappa, “Recursive state bounding by parallelotopes,” *Automatica*, vol. 32, no. 7, pp. 1049–1055, Jul. 1996.