



**HAL**  
open science

# Robust output convergence of heterogeneous networks via nonsmooth hard-threshold couplings

Félix Miranda-Villatoro

► **To cite this version:**

Félix Miranda-Villatoro. Robust output convergence of heterogeneous networks via nonsmooth hard-threshold couplings. *Nonlinear Analysis: Hybrid Systems*, 2024, 52, pp.101459. 10.1016/j.nahs.2023.101459 . hal-04333265

**HAL Id: hal-04333265**

**<https://hal.science/hal-04333265>**

Submitted on 9 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Robust output convergence of heterogeneous networks via nonsmooth hard-threshold couplings\*

Félix A. Miranda-Villatoro  
felix.miranda-villatoro@inria.fr

TRIPOP Team at Univ. Grenoble-Alpes, Inria, CNRS, LJK, Grenoble INP, 38000 Grenoble, France

## Abstract

We study a set-valued maximal monotone coupling law achieving robust output convergence in heterogeneous networks of dynamical systems with uncertainties and persistent disturbances. The coupling consists of an adaptable strategy built from normal cones to convex time-dependent sets (hard-threshold maps). To guarantee the convergence of the output mismatches to a neighborhood of the origin, only connectivity of the intrinsic graph is required (knowledge of the graph algebraic connectivity is not required), whereas only the output of the associated systems is used. Numerical simulations illustrate the effectiveness of the proposed coupling scheme.

## 1 Introduction

The study of interacting dynamical systems showing synchronized trajectories has got much attention in recent years, see *e.g.*, [5, 17, 19, 21, 22, 47] and references therein. Such intense focus has been motivated by the wide range of applications coming from almost all fields of science: power grids in engineering [36]; biological and artificial neural networks in neuroscience and computer science [32]; gene regulatory networks in biology [26]; distributed resource allocation in operations research [57]; and opinion formation of individuals in sociology [9]; are just some examples from the increasing list of applications reported in the literature. In the control and systems community, the property of synchronization has become a pivotal property, studied from different perspectives and under different contexts. At first, ideal situations were considered, where disturbances and uncertainties were absent [44, 41, 45, 52]. In such contexts, linear, and more generally, smooth approaches, (such as the master-stability function and contraction theory, see *e.g.*, [4, 50, 41]), are the dominant tools used.

Nowadays, the complexity of the studied networks has increased, as more complicated dynamics are considered together with a lack of global uniformity on the models describing the individual systems, [34, 28, 39]. Such heterogeneity in the network may be present because of different reasons. For instance, each system in the network may be described by a mathematical model that is different from that of its neighbors, with the possibility of having a state space of different dimension. In such a case, full synchronization is not possible. Nevertheless, synchronization of some of the state variables (output synchronization) may still be attainable, depending on the structure of the systems and their interactions. In other cases, even if each system belongs to the same model class, variations or uncertainties in the values of the parameters give rise to heterogeneity. Also, heterogeneity may arise due to the presence of external persistent disturbances affecting some of the agents. It is clear that when all the aforementioned sources are combined into the network, the analysis of such systems becomes a challenging task.

Some recent works dealing with the synchronization of heterogeneous networks include [12, 16, 28, 30, 34, 39, 55]. These works focus on either of the following types of interconnections: *i*) diffuse coupling; *ii*) nonlinear (high-gain) time-varying coupling; or *iii*) discontinuous coupling. For the case of diffusive coupling, it is shown in [28, 34, 39, 55] that under suitable assumptions (such as connectivity of the network,

---

\*This work is supported by the French National Research Agency in the framework of the “Investissements d’avenir” program “ANR-15-IDEX-02” and the LabEx PERSYVAL “ANR-11-LABX-0025-01”.

and semipassivity or QUAD-ness of the vector fields), the synchronization error diminishes as the amplitude of the coupling signal increases. It is shown that, in general, an infinite interconnection signal is needed in order to force the exact invariance of the synchronization manifold, so that asymptotic synchronization cannot be attained in practice. For the case of nonlinear coupling, [12, 30] showed, under similar assumptions as in the diffusive case, that a better performance is obtained with high-gain time-varying interconnections. In [30] nonlinear coupling strategies, inspired from the literature of funnel control, are studied. It is shown there that such strategies achieve high-precision in the presence of heterogeneity. However, the coupling applied requires the knowledge of the full state. Similar interaction laws were proposed in [12] regarding adaptable dead-zone maps for the case of identical systems under external noise perturbations. In such a case, the dead-zone induces an implicit funnel for the synchronization error, and the adaptation mechanism drives the funnel towards a neighborhood of the origin. Showing, once again high precision for the synchronized trajectories. Finally, for the case of discontinuous coupling, [16] established interesting properties of the network that are not shared by their smooth counterparts. Indeed, [16] showed that discontinuous couplings can achieve the asymptotic convergence of the full state with finite coupling strength. Their approach is reminiscent of sliding-mode control techniques for lumped systems, where the synchronization manifold is seen, to some extent, as a sliding surface for the synchronization error. However, it is noteworthy that the asymptotic convergence in [16] holds only at theoretical level, as the coupling is implemented in a regularized fashion to avoid the so called chattering effect. Thus, adding a boundary layer around the synchronization manifold leading to practical synchronization.

Notably, nonsmooth control techniques have shown a remarkable performance when dealing with uncertainties and disturbances, endowing the closed-loop with interesting properties such as, finite-time convergence and model reduction, see *e.g.*, [31, 33, 51]. In contrast, to achieve such distinctive performance, special attention must be put at implementation level, as without an appropriate implementation, those techniques may lead to the appearance of *chattering*, which leads to degradation in performance and the useful life of components. Regarding networks of systems, the prevalent source of nonsmoothness concerns discontinuous (switched) and impulsive couplings, these have been addressed in recent studies such as [14, 15, 24, 27, 42, 49, 54, 56]. In there, the main concern regards the synchronization of systems under interactions that change abruptly, or in cases where the communication channels put constraints on the coupling between systems. However, besides such cases, very little is known about the performance, and the implementation, of more general nonsmooth strategies. Thus, it becomes natural to consider the robust output synchronization problem under nonsmooth couplings, as a way to counteract all the effects caused by disturbances and uncertainties affecting the network.

In this paper we study the robust output convergence of heterogeneous networks using tools from the theory of differential inclusions and convex analysis. The term convergence is used in this paper, instead of the term “synchronization”, as the amplitude of the coupling signals is not constrained to be small [43, Section 1.2.1]. First, some theoretical results are presented regarding perfect output convergence in presence of uncertainties and persistent disturbances. Later, practical convergence is studied by considering a discrete-time implementation of the ideal coupling. The proposed coupling scheme has a strong connection with funnel coupling [30], adaptable dead-zone coupling [12], constrained differential inclusions [40], and perturbed Moreau’s sweeping processes [35]. Intuitively, it can be seen as a nonsmooth generalization of the funnel coupling in [30] and the dead-zone coupling in [12]. It consists in applying a correction at each time that the mismatches between the outputs of neighbors reach the boundary of a *control set*  $S(t)$ , so that  $S(t)$  is rendered positively invariant. Such correction terms are implemented via hard-thresholds to the set  $S(t)$ , so that the coupling action is inactive whenever the mismatches reside in the interior of  $S(t)$ . Additionally,  $S(t)$  is designed such that it contracts asymptotically towards zero at a rate depending on the aforementioned mismatches. In this way, output convergence is easily achieved by means of high-gain couplings. The main problem that arises with the proposed strategy concerns the well-posedness (existence and uniqueness of absolutely continuous solutions) of the interconnected systems, as the operators characterizing the hard-threshold mapping are set-valued and locally unbounded on the boundary of the set  $S(t)$ . That is, the associated differential inclusion is not of Filippov type [48, Definition 2.2]. Thus, special attention is paid to the well-posedness of the problem and sufficient conditions for the existence of solutions are presented in Theorem 6. Finally, it is shown that the proposed strategy is robust against parametric uncertainties, as well as, external disturbances. The main contributions of the paper are summarized as follows:

- The proof of existence of absolutely continuous solutions for set-valued heterogeneous networks with nonsmooth hard-threshold couplings, where the dimension of the state at each vertex is not necessarily the same.
- The robust output convergence in the presence of unmatched, persistent disturbances.
- The presentation of two numerical approaches, (one centralized and one fully distributed), for the implementation of the proposed nonsmooth schemes in digital computers.
- The extension of the funnel-coupling proposed in [30] to the multivariable, nonsmooth case.

The paper is organized as follows. The next section recalls some results from convex analysis and settles the notation used all over the paper. Section 3 formulates the problem and introduces the proposed coupling strategy. The formal proof concerning the well-posedness of the coupled system is deferred to Section 6. Section 4 studies the asymptotic properties of the interconnected system in a robust framework, whereas, Section 5 presents two specific numerical approaches for the discrete-time implementation of the proposed schemes in digital computers. The paper ends with Section 7 where conclusions are presented.

## 2 Notation and preliminaries

The Euclidean inner product is represented as  $\langle \cdot, \cdot \rangle$  and the associated norm as  $\|\cdot\|$ . The  $p$ -norm of a vector is denoted as  $\|\cdot\|_p$ , for  $p \in [1, +\infty]$ . For the sake of simplicity we drop the subindex for the 2-norm. In the cases when the argument of the norm is a matrix,  $A \in \mathbb{R}^{r \times l}$ , the induced norm is considered, that is  $\|A\| = \sup_{\|x\|=1} \|Ax\|$ . The null and range subspaces of  $A$  are represented as,  $\text{null } A$  and  $\text{rge } A$ , respectively. The matrix  $A^\dagger \in \mathbb{R}^{l \times r}$  denotes the Moore-Penrose pseudoinverse of  $A$ . The identity matrix is denoted as  $I_l \in \mathbb{R}^{l \times l}$ , whereas  $\text{Id} : \mathbb{R}^l \rightarrow \mathbb{R}^l$  denotes the identity map. The set  $\mathbb{B}_l$  denotes the closed unit ball in  $\mathbb{R}^l$  with center at zero (in the cases where the dimension is clear from the context we will drop the subindex  $l$ ). For any two nonempty sets  $\mathcal{U}, \mathcal{V} \subseteq \mathbb{R}^l$ ,  $\mathcal{U} + \mathcal{V} = \{u + v | u \in \mathcal{U}, v \in \mathcal{V}\}$ , and  $A\mathcal{U} = \bigcup_{u \in \mathcal{U}} \{Au\}$ . The interior of a set  $\mathcal{U}$  is denoted as  $\text{int } \mathcal{U}$ . The spaces  $L_1([0, T]; \mathbb{R}^l)$  and  $L_\infty([0, T]; \mathbb{R}^l)$  correspond, respectively, to the Lebesgue spaces of all absolutely integrable and essentially bounded functions, from  $[0, T]$  into  $\mathbb{R}^l$ . For a sequence  $\{f_k\}_{k \in \mathbb{N}} \subset L_1([0, T]; \mathbb{R}^l)$ , the notation  $f_k \rightharpoonup f$  denotes convergence in the weak topology of  $L_1([0, T]; \mathbb{R}^l)$ , whereas if  $\{f_k\}_{k \in \mathbb{N}} \subset L_\infty([0, T]; \mathbb{R}^l)$ , then the notation  $f_k \overset{*}{\rightharpoonup} f$  denotes convergence in the weak\* topology of  $L_\infty([0, T]; \mathbb{R}^l)$ . The reader is addressed to [10, Chapter 3] for further details on these notions of convergence.

### 2.1 Elements from convex analysis

Let  $\mathcal{U}, \mathcal{V} \subset \mathbb{R}^l$  be two nonempty, closed, convex sets. The distance from a point  $x \in \mathbb{R}^l$  to  $\mathcal{U}$  is given as

$$\text{dist}(x, \mathcal{U}) = \min_{y \in \mathcal{U}} \|x - y\| = \|x - \text{proj}(\mathcal{U}; x)\|,$$

where  $\text{proj}(\mathcal{U}; x)$  denotes the projection of  $x$  onto the set  $\mathcal{U}$ , *i.e.*,

$$\text{proj}(\mathcal{U}; x) = \arg \min_{y \in \mathcal{U}} \|x - y\|. \tag{1}$$

Let  $\varphi : \mathbb{R}^l \rightarrow \mathbb{R} \cup \{+\infty\}$  be a proper, convex, lower semicontinuous function. The convex subdifferential of  $\varphi$  at  $x$  is defined as the set-valued map

$$\partial\varphi(x) := \{w \in \mathbb{R}^l | \langle w, y - x \rangle \leq \varphi(y) - \varphi(x), \text{ for all } y \in \mathbb{R}^l\}.$$

One of the most important properties of convex subdifferentials, that is central in the developments that follow, concerns its maximal monotonicity. A set-valued map<sup>1</sup>  $\mathbf{M} : \mathbb{R}^l \rightrightarrows \mathbb{R}^l$  is monotone (in the sense of Minty-Browder) if for any two pairs  $(x_1, w_1), (x_2, w_2) \in \text{gph } \mathbf{M} := \{(x, w) \in \mathbb{R}^l \times \mathbb{R}^l | w \in \mathbf{M}(x)\}$ , the inequality

$$\langle x_1 - x_2, w_1 - w_2 \rangle \geq 0,$$

<sup>1</sup>The double arrow is used to emphasize that  $\mathbf{M}$  maps points to sets. See for instance [46, Chapter 5].

holds. In addition,  $\mathbf{M}$  is maximal monotone, if it is monotone and its graph is not strictly contained in the graph of any other monotone map. Maximal monotonicity guarantees that, for any  $r > 0$ , the map  $(\text{Id} + r\mathbf{M})^{-1}$  is single-valued and Lipschitz continuous, (indeed it is firmly nonexpansive, see [7, Definition 4.1 and Proposition 23.7]), and with domain equal to the entire space  $\mathbb{R}^l$  (Minty's theorem [7, Theorem 21.1]).

The proximal map to the function  $\varphi$  at  $x$ ,  $\text{prox}(\varphi; x)$  is the function

$$\text{prox}(\varphi; x) := \arg \min_{w \in \mathbb{R}^n} \left\{ \varphi(w) + \frac{1}{2} \|w - x\|^2 \right\}. \quad (2)$$

The proximal map to  $\varphi$  is related to the convex subdifferential of  $\varphi$  via the following expression, see *e.g.*, [7, Theorem 16.34],

$$\text{prox}(\varphi; x) = (\text{Id} + \partial\varphi)^{-1}(x). \quad (3)$$

The right-hand side of (3) is also called the *resolvent* of  $\partial\varphi$  at  $x$ , see *e.g.*, [7, Chapter 23]. It is clear from (1) and (2) that the proximal map is a generalization of the projection onto a closed, convex set, as

$$\text{proj}(\mathcal{U}; x) = \text{prox}(\psi_{\mathcal{U}}; x),$$

where  $\psi_{\mathcal{U}} : \mathbb{R}^l \rightarrow \mathbb{R} \cup \{+\infty\}$  is the indicator function of the set  $\mathcal{U}$ , such that  $\psi_{\mathcal{U}}(x) = 0$  for all  $x \in \mathcal{U}$  and  $\psi_{\mathcal{U}}(x) = +\infty$  otherwise. The set-valued map  $\mathbf{N}(\mathcal{U}; \cdot) : \mathcal{U} \rightrightarrows \mathbb{R}^l$  denotes the normal cone to  $\mathcal{U}$ , given by

$$\mathbf{N}(\mathcal{U}; x) = \partial\psi_{\mathcal{U}}(x) = \{w \in \mathbb{R}^l \mid \langle w, y - x \rangle \leq 0, \text{ for all } y \in \mathcal{U}\}.$$

Thus, it follows from (3) that

$$\text{proj}(\mathcal{U}; x) = (\text{Id} + \mathbf{N}(\mathcal{U}; \cdot))^{-1}(x) \quad (4)$$

The function  $\sigma(\mathcal{U}; u) = \sup_{p \in \mathcal{U}} \langle p, u \rangle$  denotes the support function of the set  $\mathcal{U}$  at  $u \in \mathcal{U}$ .

## 2.2 Elements from graph theory

A graph  $\mathcal{G}(\mathcal{V}, \mathcal{E})$  is a mathematical structure consisting of a set of vertices  $\mathcal{V} = \{\nu_1, \dots, \nu_N\}$  where  $\nu_i \in \mathcal{V}$  represents the  $i$ -th vertex, together with its connections, represented via a set of edges  $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ . So that, if  $\{\nu_i, \nu_j\} \in \mathcal{E}$ , then the vertices  $\nu_i$  and  $\nu_j$  are connected (*adjacent*). Throughout the paper adjacency between vertices  $\nu_i$  and  $\nu_j$  is also denoted as  $\nu_i \sim \nu_j$ . In addition, each edge  $\{\nu_i, \nu_j\} \in \mathcal{E}$  is *incident* with the vertices  $\nu_i$  and  $\nu_j$ . When the set of vertices and edges is clear from the context we will denote the graph simply as  $\mathcal{G}$ . Note that, in the set notation used,  $\{\nu_i, \nu_j\} = \{\nu_j, \nu_i\}$ , that is, the graph under consideration is *undirected*. Let  $\nu_i, \nu_j \in \mathcal{V}$ , a path of length  $m$  from  $\nu_i$  to  $\nu_j$  is a sequence of  $m + 1$  vertices  $\{\nu_k\}_{k=0}^m \subseteq \mathcal{V}$  such that  $\nu_0 = \nu_i$ ,  $\nu_m = \nu_j$ , and  $\{\nu_k, \nu_{k+1}\} \in \mathcal{E}$  for  $k \in \{0, \dots, m-1\}$ . The graph  $\mathcal{G}(\mathcal{V}, \mathcal{E})$  is *connected* if for any pair of distinct vertices  $(\nu_i, \nu_j) \in \mathcal{V} \times \mathcal{V}$  there exists a path from  $\nu_i$  to  $\nu_j$ . For each edge  $\epsilon_k = \{\nu_i, \nu_j\} \in \mathcal{E}$  a sign to each end of  $\epsilon_k$  is assigned. Such sign assignation will provide an *orientation* to the graph  $\mathcal{G}(\mathcal{V}, \mathcal{E})$ . Along all the manuscript, it is assumed that an orientation has been chosen and it is fixed. Thus, the *oriented* incidence matrix  $\Theta \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{E}|}$  is given as, see *e.g.*, [23],

$$[\Theta]_{i,k} = \begin{cases} +1, & \text{if } \nu_i \text{ is the positive end of } \epsilon_k; \\ -1, & \text{if } \nu_i \text{ is the negative end of } \epsilon_k; \\ 0, & \text{otherwise.} \end{cases}$$

If the graph  $\mathcal{G}$  has  $|\mathcal{V}|$  vertices and  $c$  connected components, then any associated incidence matrix has rank  $|\mathcal{V}| - c$ . Moreover, the graph Laplacian  $L = \Theta\Theta^\top$  and  $\Theta^\top \mathbf{1}_N = 0$ .

## 3 Problem formulation and proposed coupling strategy

Let  $\mathcal{G}(\mathcal{V}, \mathcal{E})$  be an undirected and connected graph, with vertices  $\nu_i \in \mathcal{V}$  and edges  $\{\nu_i, \nu_j\} \in \mathcal{E}$ , characterizing the connection structure between vertices. Each vertex is associated with a nonlinear system of the form,

$$\nu_i : \begin{cases} \dot{x}_i(t) = f_i(x_i(t)) + B_i u_i(t) + G_i \zeta_i(t, x_i(t)); \\ y_i(t) = C_i x_i(t), \end{cases} \quad (5)$$

where, for each  $i \in \{1, \dots, |\mathcal{V}|\}$ ,  $x_i(t) \in \mathbb{R}^{n_i}$  denotes the state of the  $i$ -th system at time  $t$ ;  $u_i(t), y_i(t) \in \mathbb{R}^m$  denote the variables available for interconnection with other systems, as indicated by the graph  $\mathcal{G}$ . The term  $\zeta_i(t, x_i(t)) \in \mathbb{R}^{p_i}$  is in general unknown, as it takes into account parametric uncertainties of the model, as well as, external disturbances affecting the  $i$ -th system. Finally, all the matrices are constant and of the appropriate dimensions, whereas each  $f_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}^{n_i}$  is assumed Lipschitz continuous.

Throughout the paper it is assumed that each function  $\zeta_i(\cdot, x_i)$  is measurable, whereas  $\zeta_i(t, \cdot)$  is a function of bounded variation, such that there exists a set-valued map  $\mathbf{Z}_i : \mathbb{R} \times \mathbb{R}^{n_i} \rightrightarrows \mathbb{R}^{p_i}$ , that is measurable in its first argument and upper-semicontinuous in its second argument, with compact and convex images and such that  $\zeta_i(t, x_i(t)) \in \mathbf{Z}_i(t, x_i(t))$  for almost all times. In addition, the following assumption on the norm of  $\zeta_i$  is considered.

**Assumption 1.** For each function  $\zeta_i$  there exist non-negative constants  $M_{i,1}$  and  $M_{i,2}$  such that the following inequality holds for all  $t \geq 0$ ,

$$\|\zeta_i(t, x(t))\| \leq M_{i,1}\|x(t)\| + M_{i,2}.$$

As stated in the introduction, we will focus on the study of a particular nonsmooth coupling law, described below, enforcing the robust agreement (against the heterogeneity of the network and external disturbances  $\zeta_i$ ) of all vertex outputs  $y_i(t)$  towards a single common trajectory or a small neighborhood of it. These two properties are formalized in the following definition.

**Definition 1.** The family of systems (5) achieves robust asymptotic output convergence if there exist coupling laws  $u_i(t)$  such that for any two indices  $i, j \in \{1, \dots, |\mathcal{V}|\}$

$$\lim_{t \rightarrow \infty} \|y_i(t) - y_j(t)\| = 0.$$

Likewise, it achieves practical output convergence if for any  $\varepsilon > 0$  there exist coupling laws  $u_i(t, \varepsilon)$ ,  $i \in \{1, \dots, |\mathcal{V}|\}$ , such that for any two indices  $i, j \in \{1, \dots, |\mathcal{V}|\}$

$$\lim_{t \rightarrow \infty} \text{dist}(y_i(t) - y_j(t); \varepsilon \mathbb{B}) = 0.$$

**Problem formulation** Given a family of systems (5) and a graph interconnection  $\mathcal{G}(\mathcal{V}, \mathcal{E})$ , our target consists on designing coupling laws  $u_i(t) = u_i(t, y_{j_1}(t), \dots, y_{j_{p_i}}(t))$ , (where each  $j_k$  is such that  $\nu_{j_k} \sim \nu_i$  and  $i \in \{1, \dots, |\mathcal{V}|\}$ ), such that robust asymptotic output convergence holds in the presence of non-vanishing disturbances.

Notice that for each vertex  $\nu_i$ , the external disturbances  $\zeta_i$  are not necessarily matched with the coupling input, that is  $\text{rge } G_i \not\subseteq \text{rge } B_i$  for some  $i \in \{1, \dots, |\mathcal{V}|\}$ . In addition, the dimension of the state space may also be different from vertex to vertex leading to a heterogeneous dynamical network. In such context, it is well-known that smooth coupling, such as diffusive coupling, cannot solve the robust convergence problem exactly, since it can provide at most practical convergence with coupling gains growing towards infinity as the mismatches between trajectories approach a neighborhood of the origin, see *e.g.*, [34, 39]. On the other hand, nonsmooth control laws are able of removing, theoretically at least, all the affections caused by uncertainties and external disturbances, whilst maintaining all control signals in a bounded region, see *e.g.* [16, 33, 51].

Motivated by the robustness properties of the set-valued hard-thresholds controllers studied in [33], it is proposed to use the following set-valued and time-dependent coupling law

$$-u_i(t) \in \sum_{j \sim i} \mathbf{N}(S_{(i,j)}(t); y_i(t) - y_j(t)), \quad (6)$$

where  $\sum_{j \sim i}$  indicates that the sum is taken over all indices  $j$  such that the vertex  $\nu_j$  is adjacent to vertex  $\nu_i$  and  $\mathbf{N}(S_{(i,j)}(t); \cdot) : S_{(i,j)}(t) \rightrightarrows \mathbb{R}^m$  denotes the normal cone to the time-dependent set  $S_{(i,j)}(t)$ . At each time instant  $t \geq 0$ , each set  $S_{(i,j)}(t) = S_{(j,i)}(t)$  in (6) is given as

$$S_{(i,j)}(t) = r_{(i,j)}(t) \mathcal{C}_{(i,j)} \subset \mathbb{R}^m, \quad (7)$$

where the sets  $\mathcal{C}_{(i,j)}$  are constant and satisfy the following standing assumption.

**Assumption 2.** Each set  $\mathcal{C}_{(i,j)} \subset \mathbb{R}^m$  is nonempty, compact, convex, symmetric (i.e.,  $\mathcal{C}_{(i,j)} = -\mathcal{C}_{(i,j)}$ ), and such that  $0 \in \text{int } \mathcal{C}_{(i,j)}$ .

Note that the symmetry condition on  $\mathcal{C}_{(i,j)}$  guarantees that  $\mathbf{N}(S_{(i,j)}(t); y_i(t) - y_j(t)) = -\mathbf{N}(S_{(j,i)}(t); y_j(t) - y_i(t))$ , so that the coupling function in (6) is undirected. Figure 2 below depicts the evolution of the graph of the normal cone over time for the case when  $\mathcal{C}_{i,j} = [-1, 1]$ . The variable  $r_{(i,j)}(t) \in \mathbb{R}$  in (7) is used to control the size of each set  $S_{(i,j)}(t)$  and it obeys

$$\dot{r}_{(i,j)}(t) = -\gamma(\|y_i(t) - y_j(t)\|)r_{(i,j)}(t), \quad (8)$$

with  $r_{(i,j)}(0) > 0$ , and  $\gamma : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a class- $\mathcal{K}$  function, that is,  $\gamma$  is strictly increasing and  $\gamma(0) = 0$ . In addition,  $\gamma$  is chosen such that there are constants  $\bar{M}_\gamma, M_\gamma \geq 0$  such that for any two outputs  $y_i(t), y_j(t)$ ,

$$\gamma(\|y_i(t) - y_j(t)\|) \leq \bar{M}_\gamma \|y_i(t) - y_j(t)\| + M_\gamma. \quad (9)$$

Note that with such initial condition,  $r_{(i,j)}(t) > 0$  for all  $t \in [0, +\infty)$  and

$$r_{(i,j)}(t) = e^{-\int_0^t \gamma(\|y_i(\tau) - y_j(\tau)\|) d\tau} r_{(i,j)}(0).$$

Thus, for each  $t \geq 0$ , each set  $S_{(i,j)}(t)$  is compact and convex, and it contracts asymptotically to the singleton  $\{0\}$  as time grows towards infinity.

The following example shows that the proposed coupling strategy (6) is not only a mathematical abstraction, but it can emerge, for instance, in the context of electrical circuits.

**Example 3.** Let us consider the nonsmooth diode network in Figure 1 carrying out the interconnection between vertices  $\nu_i$  and  $\nu_j$ .

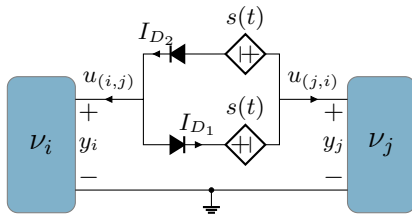


Figure 1: Ideal electrical circuit realizing the coupling (6) between systems  $\nu_i$  and  $\nu_j$ .

Let us assume that each diode satisfies an ideal complementarity condition as (see e.g., [1, Section 1.1])

$$0 \leq I_{D^k} \perp V^* - V_{D^k} \geq 0,$$

where the notation  $0 \leq P \perp Q \geq 0$  is the short form of the following three conditions: i)  $P \geq 0$ , ii)  $Q \geq 0$ , iii)  $PQ = 0$ ;  $I_{D^k}$  denotes the current flowing through the  $k$ -th diode  $D^k$ ;  $V_{D^k}$  denotes the voltage across the terminals of  $D^k$ ; and  $V^*$  denotes the threshold voltage of  $D^k$  for  $k \in \{1, 2\}$ . The ideal complementarity condition captures the behavior of the ideal diode in a simple way, since either: the current through the diode is zero and the voltage is lesser or equal to  $V^*$ , or the voltage is equal to  $V^*$  and the current is nonnegative. Setting each voltage controlled source in Figure 1 as  $s(t) = r_{(i,j)}(t) - V^*$ , it follows from Kirchhoff's laws that (in what follows the explicit dependence on time is omitted in all variables for the sake of simplicity)

$$u_{(i,j)} = -u_{(j,i)} = I_{D_2} - I_{D_1}; \quad (10a)$$

$$y_j - y_i = s + V_{D_2} = -s - V_{D_1}; \quad (10b)$$

$$0 \leq I_{D_1} \perp V^* - V_{D_1} \geq 0; \quad (10c)$$

$$0 \leq I_{D_2} \perp V^* - V_{D_2} \geq 0. \quad (10d)$$

The substitution of (10b) into the complementarity conditions (10c)-(10d) yields

$$0 \leq I_{D_1} \perp r_{(i,j)} + y_j - y_i \geq 0; \quad (11a)$$

$$0 \leq I_{D_2} \perp r_{(i,j)} + y_i - y_j \geq 0. \quad (11b)$$

It is clear from (11) that for any value of the currents  $I_{D_1}$  and  $I_{D_2}$ , the mismatch between the outputs of the  $i$ -th and  $j$ -th vertices satisfies

$$-r_{(i,j)} \leq y_j - y_i \leq r_{(i,j)}.$$

Hence,

1. If  $y_j - y_i = -r_{(i,j)}$ , then it follows from (11b) and (10a) that  $I_{D_2} = 0$  and  $u_{(i,j)} = -I_{D_1} \leq 0$ .
2. If  $y_j - y_i = r_{(i,j)}$ , then it follows from (11a) and (10a) that  $I_{D_1} = 0$  and  $u_{(i,j)} = I_{D_2} \geq 0$ .
3. If  $-r_{(i,j)} < y_j - y_i < r_{(i,j)}$ , then  $I_{D_1} = I_{D_2} = 0$  and  $u_{(i,j)} = 0$ .

Putting all cases together we arrive at the expression

$$-u_{(i,j)} \in \mathbf{N}([-r_{(i,j)}, r_{(i,j)}]; y_i - y_j).$$

Figure 2 displays the time evolution of the graph of the normal cone to the set  $r_{(i,j)}(t)[-1, 1]$  for some function  $r_{(i,j)}$  so that  $r_{(i,j)}(t) \rightarrow 0$  as  $t \uparrow \infty$ . Finally, joining all vertices of  $\mathcal{G}(\mathcal{V}, \mathcal{E})$  with coupling circuits as the one in Figure 1 produces a coupling law of the form (6) as

$$u_i = \sum_{j \sim i} u_{(i,j)}, \quad (12)$$

so that (12) indeed coincides with (6) for the scalar case where each set  $S_{(i,j)}(t) = r_{(i,j)}(t)[-1, 1]$ .

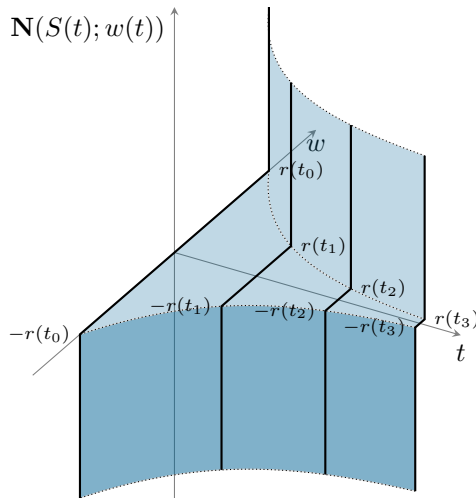


Figure 2: Time evolution of the graph of the normal cone to the time-varying set  $S(t) = [-r(t), r(t)]$ .

It is worth to emphasize that the general coupling strategy (6) is not exclusively implementable via analog electrical circuits. Indeed, in Section 5 we present two algorithms for the implementation of such coupling laws in a digital computer. Before that, the robustness of the coupled network (5)-(6) is studied in the following section.

## 4 Robust output convergence of heterogeneous networks

Let  $\Theta \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{E}|}$  be the associated oriented incidence matrix of the graph  $\mathcal{G}(\mathcal{V}, \mathcal{E})$ . The complete network with nonsmooth coupling (6) is written in compact form as

$$\dot{x}(t) \in F(x(t)) + \bar{B}u(t) + \bar{G}\mathbf{Z}(t, x(t)) \quad (13a)$$

$$\dot{r}(t) = -\Gamma((\Theta^\top \otimes I_m)y(t))r(t) \quad (13b)$$

$$y(t) = \bar{C}x(t) \quad (13c)$$

$$-u(t) \in (\Theta \otimes I_m)\mathbf{N}(S(r(t)); (\Theta^\top \otimes I_m)y(t)) \quad (13d)$$



where  $x(t) \in \mathbb{R}^n$ ,  $n = \sum_{i=1}^{|\mathcal{V}|} n_i$  is the aggregated vector of all vertex states;  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is such that  $x(t) \mapsto [f_1(x_1(t))^\top, \dots, f_{|\mathcal{V}|}(x_{|\mathcal{V}|}(t))^\top]^\top$ ;  $r(t) \in \mathbb{R}^{|\mathcal{E}|}$  contains all of the *edge* variables  $r(t) = [r_{(i_1, j_1)}(t), \dots, r_{(i_{|\mathcal{E}|}, j_{|\mathcal{E}|})}(t)]^\top$  for  $\{\nu_{i_k}, \nu_{j_k}\} \in \mathcal{E}$ ; and  $\mathbf{Z} : \mathbb{R} \times \mathbb{R}^p \rightrightarrows \mathbb{R}^p$ , where  $p = \sum_i^{|\mathcal{V}|} p_i$ , is the set-valued map  $\mathbf{Z}(t, x(t)) = \mathbf{Z}_1(t, x_1(t)) \times \dots \times \mathbf{Z}_{|\mathcal{V}|}(t, x_{|\mathcal{V}|}(t))$ . The matrices  $\bar{B} \in \mathbb{R}^{n \times |\mathcal{V}|^m}$ ,  $\bar{C} \in \mathbb{R}^{|\mathcal{V}|^m \times n}$  and  $\bar{G} \in \mathbb{R}^{n \times p}$ , are block diagonal given by  $\bar{B} = \text{Diag}(B_1, \dots, B_{|\mathcal{V}|})$ ,  $\bar{C} = \text{Diag}(C_1, \dots, C_{|\mathcal{V}|})$ , and  $\bar{G} = \text{Diag}(G_1, \dots, G_{|\mathcal{V}|})$ . The function  $\Gamma : \mathbb{R}^{|\mathcal{E}|^m} \rightarrow \mathbb{R}^{|\mathcal{E}| \times |\mathcal{E}|}$  is such that

$$\Gamma((\Theta^\top \otimes I_m)\bar{C}x(t)) = \begin{bmatrix} \gamma(\|[(\Theta^\top \otimes I_m)\bar{C}x(t)]_1\|) & & \\ & \ddots & \\ & & \gamma(\|[(\Theta^\top \otimes I_m)\bar{C}x(t)]_{|\mathcal{E}|}\|) \end{bmatrix}.$$

Finally, the set  $S(r(t))$  is such that

$$S(r(t)) = \bigtimes_{\substack{(i,j) \\ i < j \\ \{\nu_i, \nu_j\} \in \mathcal{E}}} r_{(i,j)}(t) \mathcal{C}_{(i,j)} \subset \mathbb{R}^{|\mathcal{E}|^m} \quad (14)$$

that is,  $S(r(t))$  is the Cartesian product between all the sets indexed by the edges of the graph  $\mathcal{G}$ .

We first formulate sufficient conditions for the existence of solutions of (13) and we defer the formal proof up to Section 6. By a solution of (13) we mean a pair of absolutely continuous functions  $x : \mathbb{R}_+ \rightarrow \mathbb{R}^n$  and  $r : \mathbb{R}_+ \rightarrow \mathbb{R}^{|\mathcal{E}|}$ , satisfying the inclusion (13) for almost all times  $t$ . To that end, we impose the following assumption.

**Assumption 4.** *There exists a non-singular matrix  $R \in \mathbb{R}^{n \times n}$  such that*

$$R\bar{B} = R^{-\top}\bar{C}^\top. \quad (15)$$

Assumption 4 constraints the *relative degree* (with respect to the input-output pair  $(u(t), y(t))$ ) of the network (13) to  $\{1, \dots, 1\}$ , which is a necessary condition for passivity of the connected system (13), see e.g., [11]. However, since at this point the only assumption we make on the vector fields  $f_i$  concerns their Lipschitz continuity, the composed network is not necessarily passive. This lack of passivity allows us to consider the robust agreement of systems with interesting behaviors, such as self-sustained oscillations or chaotic solutions. It is easy to see that if for each  $i \in \{1, \dots, |\mathcal{V}|\}$  there exists a symmetric, positive definite matrix,  $P_i \in \mathbb{R}^{n_i \times n_i}$ , such that  $P_i B_i = C_i^\top$ , then  $R = \text{Diag}(P_1^{1/2}, \dots, P_{|\mathcal{V}|}^{1/2})$  satisfies Assumption 4. Notice that in a more general setting we may have some systems for which there is no symmetric, positive definite matrix,  $P_i$ , satisfying  $P_i B_i = C_i^\top$ . Nevertheless, it might still exist  $R \in \mathbb{R}^n$  satisfying (15). In addition to Assumption 4 we also consider the following standing assumption.

**Assumption 5.** *The set of admissible initial conditions is compact so that there is a constant  $\bar{R} > 0$ , such that  $\|r(0)\| \leq \bar{R}$ , and*

$$(\Theta^\top \otimes I_m)y(0) \in S(r(0)).$$

Assumption 5 is necessary for the existence of absolutely continuous solutions. Otherwise, a state jump at  $t = 0$  will take place so that  $(\Theta^\top \otimes I_m)y(0^+) = \lim_{t \rightarrow 0^+} (\Theta^\top \otimes I_m)y(t) \in S(r(0^+))$ .

We are now ready to state the main result concerning the existence of solutions of (35).

**Theorem 6.** *Let Assumptions 1, 2, 4 and 5 hold. Then, there exist absolutely continuous functions  $x : [0, T] \rightarrow \mathbb{R}^n$  and  $r : [0, T] \rightarrow \mathbb{R}^{|\mathcal{E}|}$  that satisfy (13) almost everywhere in  $[0, T]$ .*

The proof of Theorem 6 is postponed until Section 6. We now change focus to the consequences of Theorem 6 concerning the robust output convergence problem.

*Remark 1.* Note that in Theorem 6 the final time  $T$  is arbitrary but finite, and some upper-bounds in the proof depend on such  $T$ . In terms of the behavior, this translates into that fact that, even though there are not finite escape times, it is possible for the trajectories to grow unbounded as time evolves. An alternative approach consists in considering extra conditions, regarding the vertex dynamics, in order to guarantee the existence of a compact, positively invariant region, having in this way an uniform upper-bound on the state variables. For instance, if each vertex is assumed semipassive [44], then the trajectories of (13) are guaranteed to be uniformly bounded in the whole domain  $[0, +\infty)$ . Namely, the  $i$ -th vertex is semipassive if there exist  $\rho > 0$  and a continuously differentiable, radially unbounded function,  $V_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}_+$  such that for any admissible input  $u_i(t)$ ,

$$\frac{d}{dt}V(x_i(t)) \leq y_i(t)^\top u_i(t) - H_i(x_i(t))$$

where  $H_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}$  is such that  $H_i(x_i(t)) \geq M > 0$  whenever  $\|x_i(t)\| > \rho$ . Thus, the derivative of the positive definite function  $V(x) = \sum_{i=1}^{|\mathcal{V}|} V_i(x_i)$  is

$$\begin{aligned} \frac{d}{dt}V(x) &\leq \sum_{i=1}^{|\mathcal{V}|} y_i^\top u_i - \sum_{i=1}^{|\mathcal{V}|} H_i(x_i) \\ &\leq -\sum_{i=1}^{|\mathcal{V}|} \sum_{j \sim i} y_i^\top \mathbf{N}(S(r_{(i,j)}(t)); y_i - y_j) - \sum_{i=1}^{|\mathcal{V}|} H_i(x_i) \\ &= -y^\top (\Theta \otimes I_m) \mathbf{N}(S(r(t)); (\Theta^\top \otimes I_m)y) - \sum_{i=1}^{|\mathcal{V}|} H_i(x_i) \\ &\leq -\sum_{i=1}^{|\mathcal{V}|} H_i(x_i), \end{aligned} \tag{16}$$

where the last inequality above follows from the maximal monotonicity of the normal cone map. Since the right-hand side of (16) is strictly negative whenever  $\|x_i\| > \rho$  for some  $i \in \{1, \dots, |\mathcal{V}|\}$ , and the function  $V$  is radially unbounded, then all trajectories converge to the largest invariant region contained in the compact set

$$\Omega_L = \{x \in \mathbb{R}^n \mid V(x) \leq L\},$$

where  $L$  is such that  $\sqrt{|\mathcal{V}|}\rho\mathbb{B} \subset \Omega_L$ , and boundedness of trajectories of (13) follows.

The well-posedness result in Theorem 6 conveys important consequences regarding the output convergence of general heterogeneous networks in which each agent may be subject to external persistent disturbances. Namely, well-posedness of the network (13) implies that  $(\Theta^\top \otimes I_m)y(t) \in S(r(t))$  for all times  $t \geq 0$ . Moreover, it follows from the dynamics of  $r(t)$  in (13) together with the strict positivity of  $\Gamma$  that  $r(t) \rightarrow 0$  as  $t \rightarrow \infty$ . Therefore,

$$(\Theta^\top \otimes I_m)y(t) \rightarrow 0 \text{ as } t \rightarrow \infty.$$

The following corollary is thus an immediate consequence of Theorem 6.

**Corollary 7.** *Under the assumptions of Theorem 6, the coupling law (6) achieves the robust asymptotic output convergence of the family of systems (5).*

Note that, in the general heterogeneous case, the full convergence of the state is not guaranteed since, in principle, the dimension of the state of each individual system might differ from that of its neighbors. Nevertheless, in the cases when all individual systems have the same state dimension, then extra conditions can guarantee full convergence.

**Definition 2.** The network (13) is asymptotically zero-state detectable if for any two systems  $\nu_a, \nu_b$  such that the output mismatch  $y_a - y_b \rightarrow 0$ , whereas the terms  $\tilde{u}_s(t) = B_s u_s(t) + G_s \zeta(t, x_s(t))$ ,  $s \in \{a, b\}$  satisfy,  $\tilde{u}_a(t) \rightarrow 0$ ,  $\tilde{u}_b(t) \rightarrow 0$  implies  $x_a(t) - x_b(t) \rightarrow 0$ .

**Corollary 8.** *Let all assumptions of Theorem 6 hold. If in addition the dimension of the state is the same for each individual system and the incremental dynamics of the network is asymptotically zero-state detectable, then the coupling (6) achieves the full-state convergence of the family of systems (5).*

The results concerning asymptotic output convergence of the network are valid only when the ideal set-valued map in (6) is used. However, in practice, it is not possible to have a real-world implementation of the ideal coupling as depicted in Section 3, since it requires the use of ideal components. In real-life experiments there are losses due to parasitic resistance effects and unmodeled dynamics. In the next section an implementable coupling law is proposed via the implicit discretization of the continuous-time coupling (6), so that the convergence of the output towards a unique trajectory is maintained with a precision depending on the sampling step time  $h$  (practical output convergence).

## 5 Numerical implementation of the coupling strategy

Nowadays, due to the great performance and accessibility of digital electronic devices, it may be useful to implement the coupling law discussed above using a discrete-time scheme in digital microcomputers. To make such digital implementation, it becomes necessary to pay extra attention to the discretization used, as it is well known, explicit discretization schemes are prone to the appearance of chattering, which leads to degradation of closed-loop performance or even loss of stability, see *e.g.*, [1, 6]. In this section we study two approaches for computing a discrete-time coupling law based on (6).

### 5.1 Implementation via implicit discretization

The discretization considered in this subsection is based on the discretization scheme presented in [2], where the set-valued component is discretized implicitly and the disturbances are ignored for the selection process. It is similar to the discretization used in the proof of Theorem 6 (see Section 6.2), but in the present context, we are interested only in the output dynamics, as the full state is assumed unknown. Concretely, in order to compute a selection of the generalized equation (6) we consider the output dynamics

$$\dot{y}(t) = \bar{C} (F(x(t)) + \bar{G}\zeta(t, x(t))) + \bar{C}\bar{B}u(t) \quad (17a)$$

$$\dot{r}(t) = -\Gamma ((\Theta^\top \otimes I_m)y(t)) r(t) \quad (17b)$$

$$-u(t) \in (\Theta \otimes I_m)\mathbf{N}(S(r(t)); (\Theta^\top \otimes I_m)y(t)) , \quad (17c)$$

where  $x(t)$  is the state of the network (13) at time  $t \geq 0$ . The discretized output dynamics is

$$y^{k+1} = \tilde{y}^{k+1} + h\bar{C}(F(x^k) - F(\bar{C}^\dagger y^k) + \bar{G}\zeta^k) \quad (18a)$$

$$\tilde{y}^{k+1} = y^k + h\bar{C}F(\bar{C}^\dagger y^k) + h\bar{C}\bar{B}u^{k+1} \quad (18b)$$

$$r^{k+1} = (I + h\Gamma ((\Theta^\top \otimes I_m)y^k))^{-1} r^k \quad (18c)$$

$$-u^{k+1} \in (\Theta \otimes I_m)\mathbf{N}(S(r^{k+1}); (\Theta^\top \otimes I_m)\tilde{y}^{k+1}) . \quad (18d)$$

The new variable  $\tilde{y}^{k+1}$  in (18) plays the role of a *nominal output*. Its role consists in making the selection strategy independent of unknown data, (*i.e.*, independent of  $\bar{C}(F(x^k) + \bar{G}\zeta^k)$ ). Note that, in the cases when  $\bar{C}$  is invertible, as for instance  $\bar{C} = I_n$ , the difference  $F(x^k) - F(\bar{C}^\dagger y^k) = 0$  for all  $x^k$  and all  $y^k$ . It is also noteworthy to recall that, in general, the complete state is not available, so that the term  $F(x^k)$  is unknown. A possible way to reduce the level of uncertainty, consists in the design of state-observers as is done in [3, 37]. In this work the coupling strategy is static, so that no observer design is discussed. This, in order to keep the numerical implementation of the coupling law as simple as possible.

The following corollary establishes the well-posedness of the *implicitly* defined coupling (18b)-(18d), as well as, its effectiveness for solving the practical output convergence problem in a discrete-time context.

**Corollary 9.** *Let all assumptions of Theorem 6 hold. For the closed-loop heterogeneous network (18) the coupling action is given explicitly as*

$$u^{k+1} = -\frac{1}{h}\Lambda^{-1} (\text{Id} - \text{proj}(L^\dagger S(r^{k+1}); \bullet)) (L^\dagger \Lambda^{-1}(y^k + h\bar{C}F(\bar{C}^\dagger y^k))), \quad (19)$$

where  $L = (\Theta^\top \otimes I_m)\Lambda$  and  $\Lambda = (\bar{C}\bar{B})^{\frac{1}{2}}$ . Moreover, the “nominal output”  $\tilde{y}^{k+1} \in S(r^{k+1})$  for all  $k \geq 0$  and if the function  $F$  and  $\zeta^k$  are uniformly bounded then the network (18)-(19) achieves practical output convergence.

*Proof.* It follows from Assumption 4 that the product  $\bar{C}\bar{B}$  is symmetric and positive definite. Let  $\Lambda = \Lambda^\top \succ 0$  be such that  $\Lambda^2 = \bar{C}\bar{B}$ . Then, the change of variables  $w^k = \Lambda^{-1}y^k$ , yields,

$$w^{k+1} = \tilde{w}^{k+1} + h\Lambda^{-1}\varphi^k \quad (20a)$$

$$\tilde{w}^{k+1} = w^k + hp^k + hv^{k+1} \quad (20b)$$

$$r^{k+1} = \tilde{\Gamma}(w^k)r^k \quad (20c)$$

$$-v^{k+1} \in L^\top \mathbf{N}(S(r^{k+1}); L\tilde{w}^{k+1}) \quad (20d)$$

where  $v^{k+1} = \Lambda u^{k+1}$ ,  $\varphi^k = \bar{C}(F(x^k) - F(\bar{C}^\dagger \Lambda w^k) + \bar{G}\zeta^k)$ ,  $p^k = \Lambda^{-1}\bar{C}F(\bar{C}^\dagger \Lambda w^k)$ ,  $\tilde{\Gamma}(w^k) = (I + h\Gamma(Lw^k))^{-1}$  and  $L = (\Theta^\top \otimes I_m)\Lambda$ . Note that  $p^k$  is assumed to be known, as it depends only on nominal parameters and the measurable output  $y^k = \Lambda w^k$ . Setting  $w_\perp^k = L^\dagger L w^k$  and  $w_\parallel^k = (I - L^\dagger L)w^k$  as the projections onto  $(\text{null } L)^\perp$  and  $\text{null } L$  respectively, (similarly for  $p_\perp^k$  and  $p_\parallel^k$ ), it follows that

$$w^{k+1} = \tilde{w}^{k+1} + h\Lambda^{-1}\varphi^k \quad (21a)$$

$$\tilde{w}_\perp^{k+1} = w_\perp^k + hp_\perp^k + hv^{k+1} \quad (21b)$$

$$\tilde{w}_\parallel^{k+1} = w_\parallel^k + hp_\parallel^k \quad (21c)$$

$$r^{k+1} = \tilde{\Gamma}(w_\perp^k)r^k \quad (21d)$$

$$-v^{k+1} \in \mathbf{N}(\hat{S}(r^{k+1}); \tilde{w}_\perp^{k+1}) \quad (21e)$$

where  $\hat{S}(r^{k+1}) = \{s \in \mathbb{R}^{|\mathcal{V}|m} | Ls \in S(r^{k+1})\} = \{s \in \mathbb{R}^{|\mathcal{V}|m} | s \in L^\dagger S(r^{k+1})\}$ . It follows from (21b) and (21e) that

$$w_\perp^k + hp_\perp^k - \tilde{w}_\perp^{k+1} \in h\mathbf{N}(\hat{S}(r^{k+1}); \tilde{w}_\perp^{k+1})$$

Thus, it follows from [7, Theorem 3.14] that

$$\tilde{w}_\perp^{k+1} = \text{proj}\left(\hat{S}(r^{k+1}); w_\perp^k + hp_\perp^k\right), \quad (22)$$

and the substitution of (22) back into (21b) gives us an explicit expression for  $v^{k+1}$  as

$$v^{k+1} = -\frac{1}{h} \left( \text{Id} - \text{proj}(\hat{S}(r^{k+1}); \cdot) \right) (w_\perp^k + hp_\perp^k),$$

from where (19) follows. Also note that (22) implies that  $\tilde{w}_\perp^{k+1} \in \hat{S}(r^{k+1})$ , that is

$$L\tilde{w}_\perp^{k+1} = L\tilde{w}^{k+1} = (\Theta^\top \otimes I_m)\tilde{y}^{k+1} \in S(r^{k+1}),$$

and (18a) leads us to

$$(\Theta^\top \otimes I_m)(y^{k+1} - h\bar{C}(F(x^k) - F(\bar{C}^\dagger y^k) + \bar{G}\zeta^k)) \in S(r^{k+1}). \quad (23)$$

If the map  $F$  and  $\zeta^k$ , are uniformly bounded then there is a finite  $M > 0$  such that

$$(\Theta^\top \otimes I_m)y^{k+1} \in S(r^{k+1}) + hM\mathbb{B}.$$

So that the error is dependent of the sampling time  $h$ . Finally, as (21d) is Schur stable, then practical output convergence follows. This concludes the proof.  $\square$

*Remark 2.* The assumption regarding the uniform boundedness of  $F$  and  $\zeta$ , is sufficient but not necessary. For instance, if  $F$  is not uniformly bounded but Lipschitz continuous instead, then it follows from (23) that

$$(\Theta^\top \otimes I_m)y^{k+1} \in S(r^{k+1}) + h(M_1\|x^k\| + M_2)\mathbb{B},$$

for some finite  $M_1, M_2 \geq 0$ , so that the outputs remain close to each other whenever the trajectories rest in an invariant compact region and the sampling time  $h$  is small enough. Further assumptions on the individual vector fields, such as, semipassivity [44], incremental dissipativeness [45], or QUAD-ness of the vector field [18], are common to ensure the existence of a compact invariant region of the state, allowing to relax the uniform boundedness of both  $F$  and  $\zeta$ , see Remark 1 in Section 4.

It is worth to point out that the implicit discretization scheme in (18) is centralized, as it uses the pseudo-inverse of  $L$ , so that in order to compute the individual input  $u^{k+1}$  to each agent, it is necessary to have access to the output of the entire network at each time step. The following subsection presents a simple way of generating a distributed approach in order to achieve practical output convergence of the network, via a simple regularization of the ideal coupling.

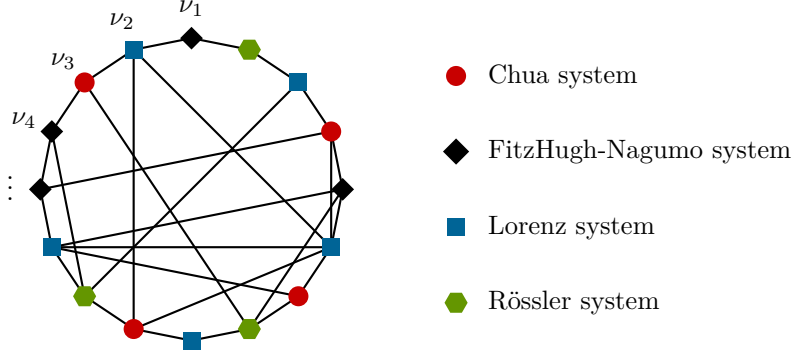


Figure 3: Heterogeneous network of systems. At each time the input to the  $i$ -th vertex is computed from the outputs of its neighbors. The parameters of each vertex system are different, even if they belong to the same class.

**Example 10.** Let us consider an heterogenous small-world network [38] consisting of 16 vertices, where each vertex is of one of the following four different type of systems: A) Chua chaotic system, B) FitzHugh-Nagumo oscillator, C) Lorentz chaotic system, and D) Rössler chaotic system. Figure 3 displays the network configuration.

Each vertex is of the form

$$\begin{aligned}\dot{x}_i(t) &= (A_i + \Delta A_i)x_i(t) + (E_i + \Delta E_i)h_i(t) + B_i u_i(t) \\ g_i(t) &= (F_i + \Delta F_i)x_i(t) \\ y_i(t) &= C_i x_i(t)\end{aligned}$$

Let  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ , be the sets of indices for systems of type A), B), C), and D) respectively. The set of nominal parameters,  $A_i, B_i, C_i, E_i$  and  $F_i$ , is shown in Table 1 for each class of vertex system  $i \in \mathcal{A} \cup \mathcal{B} \cup \mathcal{C} \cup \mathcal{D}$ . Note that the dimension of the internal states may vary from vertex to vertex. So that in this case full convergence is not possible to achieve. The terms  $\Delta A_i, \Delta E_i, \Delta F_i$  denote parametric uncertainties for each model class. Each uncertainty term is taken as a sample from a standard distribution, of appropriate dimensions, with zero mean and standard deviation  $\sigma = 0.1$ . The function  $\gamma : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  used for setting the rate of descend of  $r^k$  is set as,

$$\gamma(\eta) = \begin{cases} \eta, & \text{if } \eta > \delta; \\ 0, & \text{otherwise.} \end{cases}$$

where  $\delta = 0.5$ . In this setting, the input to each vertex is computed in a centralized fashion via (19). Figure 4 shows the state  $x_i(t)$ , the output  $y_i(t)$ , and the input  $u_i(t)$  of each vertex in the network with a sampling time of  $h = 10\text{ms}$ .

## 5.2 Implementation via regularization

The coupling law (6) is largely inspired by the electrical network in Figure 1. Thus, with the end of obtaining a distributed numerical scheme we consider a regularized version of the ideal electrical coupling discussed in Example 3. Specifically, let us consider the circuit shown in Figure 5. After simple computations, similar to those shown in Example 3, we obtain that the regularized coupling circuit obeys

$$-u_{(i,j)} = u_{(j,i)} \in \mathbf{N}(S_{(i,j)}, y_i - y_j + Ru_{(i,j)}). \quad (24)$$

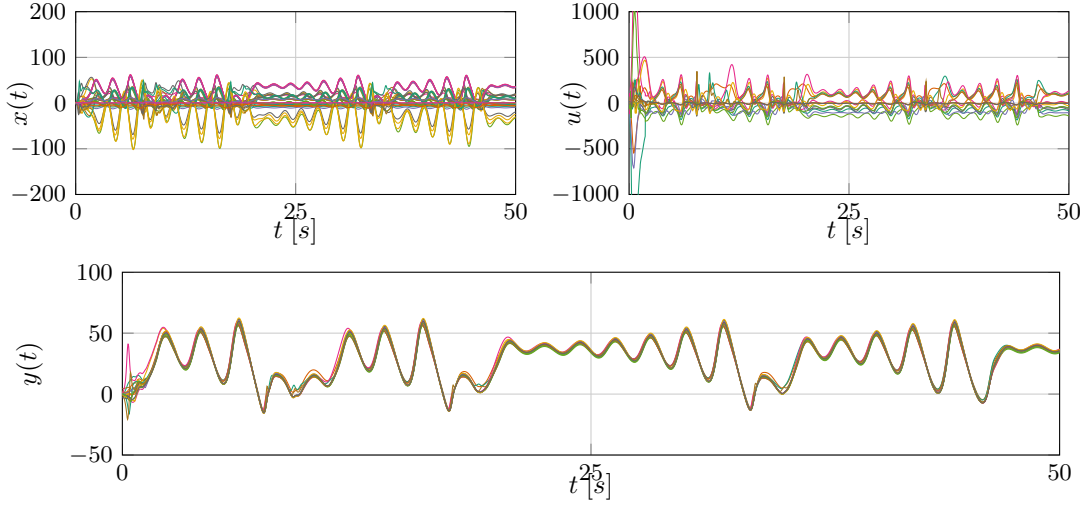


Figure 4: (Upper left) Time trajectories of states of vertex systems  $\nu_i$ . (Upper right) Time trajectories of coupling inputs  $u_i$  for each vertex. (Bottom) Time trajectories of output signals of each vertex system. The coupling is given by (19).

	$A_i$	$B_i$	$C_i$	$E_i$	$F_i$	$h_i$
$i \in \mathcal{A}$	$\begin{bmatrix} -8.8 & 8.8 & 0 \\ 1 & -1 & 1 \\ 0 & -15 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$	$B_i^\top$	$\begin{bmatrix} 8.8 \\ 0 \\ 0 \end{bmatrix}$	$B_i^\top$	$\tanh(2g_i) + 0.7g_i$
$i \in \mathcal{B}$	$\begin{bmatrix} 0 & -2 \\ 20 & 2 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$B_i^\top$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$E_i^\top$	$-g_i^3 + 5.5g_i$
$i \in \mathcal{C}$	$\begin{bmatrix} -10 & 10 & 0 \\ 28 & -1 & 0 \\ 0 & 0 & -\frac{8}{3} \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$B_i^\top$	$\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$	—	$\begin{bmatrix} -x_1x_3 \\ x_1x_2 \end{bmatrix}$
$i \in \mathcal{D}$	$\begin{bmatrix} -0.05 & -0.5 & -1 \\ 1 & 0.133 & 0 \\ 0 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}^\top$	$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^\top$	$\max\{0, 15(g_i - 3)\}$

Table 1: Nominal parameters for the three classes of systems for the network in Figure 3.

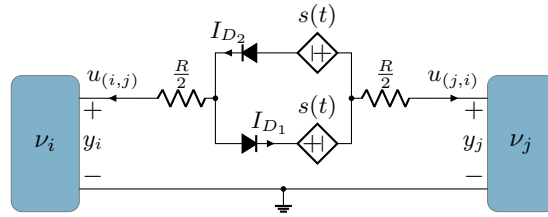


Figure 5: Regularized coupling circuit

Recalling that  $\mathbf{N}(S, y) = \partial\psi_S(y)$  is a convex cone it follows that for any  $R > 0$

$$-Ru_{(i,j)} \in \mathbf{N}(S_{(i,j)}; y_i - y_j + Ru_{(i,j)}). \quad (25)$$

Hence,

$$y_i - y_j \in (\text{Id} + \mathbf{N}(S_{(i,j)}; \bullet))(y_i - y_j + Ru_{(i,j)}),$$

and from (4) we retrieve the explicit expression for  $u_{(i,j)}$  as

$$u_{(j,i)} = -u_{(i,j)} = \frac{1}{R} (\text{Id} - \text{proj}(S_{(i,j)}; \bullet)) (y_i - y_j). \quad (26)$$

The coupling (26) is single-valued and it is shown below that it achieves practical output convergence for values of  $R$  sufficiently small. Figure 6 depicts the coupling function (26) for the scalar case.

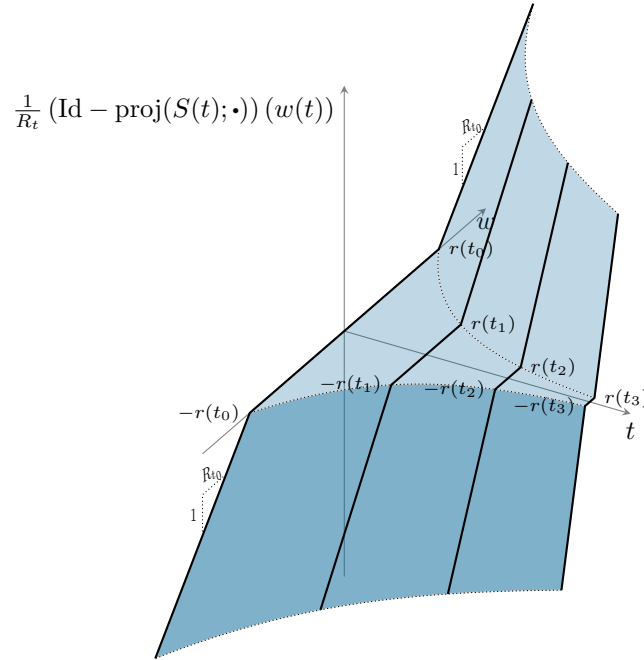


Figure 6: Time evolution of the graph of the coupling law in (26). In this case as  $t \uparrow \infty$ ,  $R_t \downarrow 0$  and the graph of the coupling map approaches the graph of the normal cone to the set  $S(t)$ , see Figure 2.

Note that, (26) is independent of the parameters of the vertices dynamics, as only the associated outputs and the set  $S_{(i,j)}$  are needed. Thus, (26) can be implemented in a distributed fashion. It is also important to remark that (26) is a Lipschitz continuous function of the output mismatch  $y_i - y_j$  and therefore the network (13)–(26) does not have finite-escape times, that is, it is well-posed in the entire domain  $[0, +\infty)$ . Moreover, in the limit as  $R \downarrow 0$ , the network is also well-posed, as in that case, the coupling (24) coincides with the original coupling (6) via (12), see Theorem 6 and the remark after it.

**Corollary 11.** *Under the assumptions of Theorem 6, the regularized coupling*

$$u_i^{k+1} = \sum_{j \sim i} u_{(i,j)}^{k+1}, \quad (27)$$

where  $u_{(i,j)}^{k+1}$  is given by

$$-u_{(i,j)}^{k+1} = u_{(j,i)}^{k+1} = \frac{1}{R_{(i,j)}^k} (\text{Id} - \text{proj}(S(r_{(i,j)}^{k+1}); \bullet)) (y_i^k - y_j^k), \quad (28)$$

and the regularization parameters  $R_{(i,j)}^k \in \mathbb{R}^+$  satisfy

$$R_{(i,j)}^k = R_{(j,i)}^k := \min \left\{ \frac{h\delta_1}{\|u_{(i,j)}^k\|}, \delta_2 \right\}, \quad (29)$$

for some  $\delta_i > 0$   $i \in \{1, 2\}$ , achieves the practical output convergence of (13) whenever the time step  $h$  is sufficiently small.

*Proof.* It follows from (28) that

$$y_i^k - y_j^k + R_{(i,j)}^k u_{(i,j)}^{k+1} = \text{proj}(S(r_{(i,j)}^{k+1}); y_i^k - y_j^k).$$

Hence,  $\|R_{(i,j)}^k u_{(i,j)}^{k+1}\| = \text{dist}(y_i^k - y_j^k; S(r_{(i,j)}^{k+1}))$ . Consequently,

$$y_i^k - y_j^k \in S(r_{(i,j)}^{k+1}) + R_{(i,j)}^k \|u_{(i,j)}^{k+1}\| \mathbb{B} \subset S(r_{(i,j)}^{k+1}) + h\delta_1 \frac{\|u_{(i,j)}^{k+1}\|}{\|u_{(i,j)}^k\|} \mathbb{B}.$$

In the limit as  $h \downarrow 0$ ,  $\frac{\|u_{(i,j)}^{k+1}\|}{\|u_{(i,j)}^k\|} \rightarrow 1$ , (notice that such limit is well-defined, since in the limit we retrieve the strategy (6) whose well-posedness is guaranteed by Theorem 6), so that for  $h$  sufficiently small the output mismatch  $y_i^k - y_j^k$  lies in a neighborhood of  $S(r_{(i,j)}^{k+1})$  and practical output convergence follows. This concludes the proof.  $\square$

**Example 12.** Let us consider again the heterogeneous network of systems described in Example 10 above. This time each coupling is implemented in a distributed fashion using (28). The nominal parameters of each vertex are the same as shown in Table 1 and the function  $\gamma$  describing the rate of descend of the variables  $r_{(i,j)}$  also remains unchanged. The disturbances are taken in the same manner as in Example 10. In this example, the sampling time is decreased to  $h = 2.5\text{ms}$ , (as for a sampling time of  $h = 10\text{ms}$ , the coupling (28) fails to achieve the ultimate boundedness of the mismatch  $y_i^k - y_j^k$ ). Note that, in this case, no knowledge regarding the vector field  $F$  is used for the computation of the coupling signals. The regularization parameters  $R_{(i,j)}^k$  are set as in (29) with  $h\delta_1 = 25$  and  $\delta_2 = 0.1$ . Figure 7 shows the time trajectories of the state, control input, and output of each vertex system, illustrating the practical output convergence of the heterogeneous network of Figure 3.

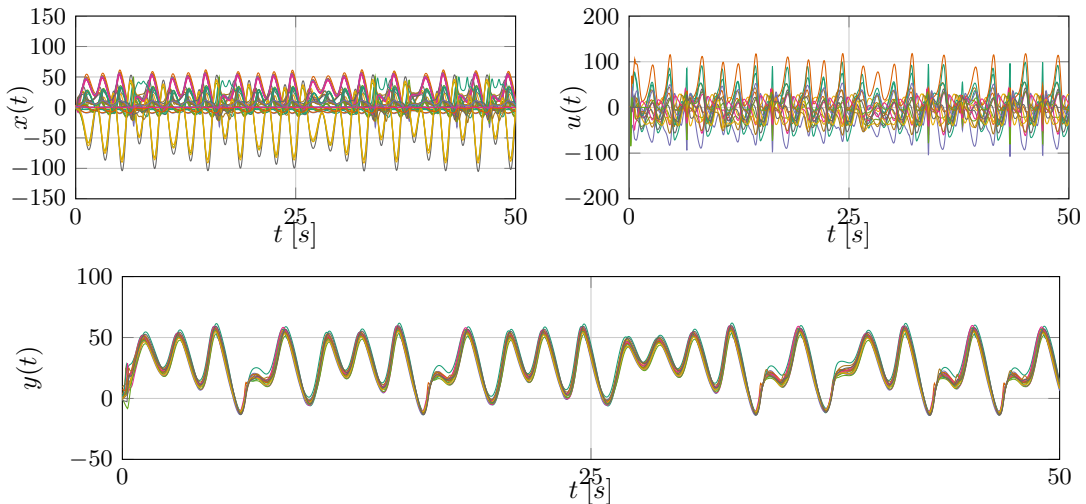


Figure 7: (Upper left) Time trajectories of states of vertex systems  $\nu_i$ . (Upper right) Time trajectories of coupling inputs  $u_i$  for each vertex. (Bottom) Time trajectories of output signals for each vertex.

Finally, Figure 8 displays the sum of square error signals for the coupling schemes (19) and (28) under the same sampling rate. As expected, there is a trade-off between centrality and precision, since the centralized scheme (19) leads to a lower error compared to the fully distributed scheme (28).



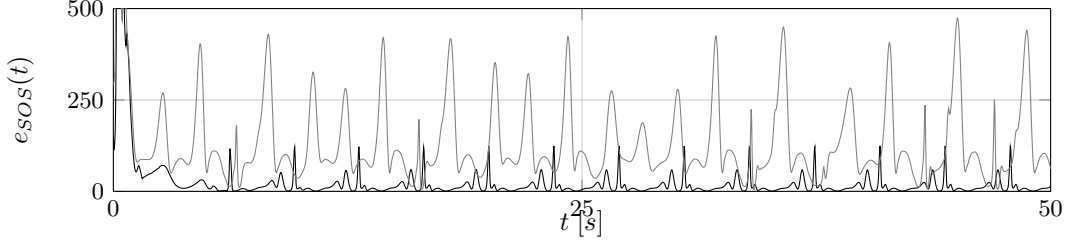


Figure 8: Time evolution of the sum of squares error signal  $e_{SOS}(t) := \frac{1}{2} \sum_{i=1}^{|\mathcal{V}|} \sum_{j \sim i} \|y_i(t_k) - y_j(t_k)\|^2$  for coupling schemes (19)(black) and (28) (gray) with a common sampling period of  $h = 2.5ms$ .

## 6 Existence of solutions of the coupled network

This section presents the proof of Theorem 6 regarding existence of absolutely continuous solutions of the differential inclusion (13). The proof is an adaptation and an extension of the proof in [25] and is presented here for completeness. The proof is divided in several subsections. First, a change of coordinates and an useful decomposition of the dynamics is performed, putting the model into a more suitable structure for its analysis. Second, a sequence of approximate solutions is constructed based on its associated discrete-time model. Third, it is shown that these sequences of approximate solutions converge in appropriated spaces. Finally, it is shown that the limit functions are indeed trajectories of the continuous-time model.

### 6.1 State transformation and model decomposition

Let us consider the following change of variables

$$z(t) = Rx(t), \quad (30)$$

where  $R$  satisfies Assumption 4. Thus, (13) is transformed into

$$\begin{cases} \dot{z}(t) \in \tilde{F}(z(t)) - \tilde{\Theta}^\top \mathbf{N}(S(r(t)); \tilde{\Theta}z(t)) + (I_N \otimes RG)\mathbf{Z}(t, (I_{|\mathcal{V}|} \otimes R^{-1})z(t)), \\ \dot{r}(t) = -\Gamma(\tilde{\Theta}z(t))r(t), \end{cases} \quad (31)$$

where,  $\tilde{F} = R \circ F \circ R^{-1}$  and  $\tilde{\Theta} = (\Theta^\top \otimes I_m)\bar{C}R^{-1} = (\Theta^\top \otimes I_m)\bar{B}^\top R^\top \in \mathbb{R}^{|\mathcal{E}|m \times n}$ . It follows from Assumption 2 that  $0 \in \text{int } S(r(t))$  for all  $t \geq 0$ . Consequently, for any  $\vartheta \in \text{null } \tilde{\Theta}$ ,  $\tilde{\Theta}\vartheta \in S(r(t))$  for all  $t \geq 0$ . Hence,  $\text{rge}(\tilde{\Theta}) \cap \text{int } S(r(t)) \neq \emptyset$  for all  $t$ , and it follows from [7, Proposition 16.32] that

$$\tilde{\Theta}^\top \mathbf{N}(S(r(t)); \tilde{\Theta}z(t)) = \partial(\psi_{S(r(t))} \circ \tilde{\Theta})(z(t)).$$

It follows easily from the definition of normal cone and the chain rule for convex subdifferentials that

$$\tilde{\Theta}^\top \mathbf{N}(S(r(t)); \tilde{\Theta}z(t)) = \mathbf{N}(\tilde{S}(r(t)); z(t)), \quad (32)$$

where

$$\tilde{S}(r(t)) = \left\{ \vartheta \in \mathbb{R}^n \mid \tilde{\Theta}\vartheta \in S(r(t)) \right\} \subset \mathbb{R}^n, \quad (33)$$

is closed and convex. Hence, (31) transforms into

$$\begin{cases} \dot{z}(t) \in \tilde{F}(z(t)) - \mathbf{N}(\tilde{S}(r(t)); z(t)) + \tilde{\mathbf{Z}}(t, z(t)) \\ \dot{r}(t) = -\Gamma(\tilde{\Theta}z(t))r(t) \end{cases} \quad (34)$$

where  $\tilde{\mathbf{Z}}(t, z(t)) = R\bar{G}\mathbf{Z}(t, R^{-1}z(t))$ .

In what follows we denote the orthogonal projection of any  $\eta \in \mathbb{R}^n$  onto the subspace  $\text{null } \tilde{\Theta}$  as  $\eta_\parallel$ , and the complementary projection onto  $(\text{null } \tilde{\Theta})^\perp$  as  $\eta_\perp$ . That is, let  $\Pi$  be the matrix representation of the projection onto  $\text{null } \tilde{\Theta}$ . Then, for any  $\eta \in \mathbb{R}^n$ ,  $\eta = \eta_\perp + \eta_\parallel$ , where  $\eta_\perp = (I - \Pi)\eta$  and  $\eta_\parallel = \Pi\eta$ .

Hence, letting  $z(t) = z_{\perp}(t) + z_{\parallel}(t)$ , (34) is rewritten as,

$$\dot{z}_{\perp}(t) \in \tilde{F}_{\perp}(z(t)) - \mathbf{N}(\tilde{S}(r(t)); z_{\perp}(t)) + \tilde{\mathbf{Z}}_{\perp}(t, z(t)) \quad (35a)$$

$$\dot{z}_{\parallel}(t) \in \tilde{F}_{\parallel}(z(t)) + \tilde{\mathbf{Z}}_{\parallel}(t, z(t)) \quad (35b)$$

$$\dot{r}(t) = -\Gamma(\tilde{\Theta}z_{\perp}(t))r(t) \quad (35c)$$

where we have used the fact that  $\Pi\tilde{\Theta}^{\top} = 0$ , for  $\Pi$  a projector onto null  $\tilde{\Theta}$ , together with (32).

## 6.2 Construction of approximate solutions

Clearly, that well-posedness of (13) is equivalent to that of (35). Thus, we continue by discretizing in time the dynamics (35) using a semi-implicit Euler method.

$$z_{\perp}^{k+1} \in z_{\perp}^k + h\tilde{q}_{\perp}(t^k, z^k) - h\mathbf{N}(\tilde{S}(r^{k+1}); z_{\perp}^{k+1}) \quad (36a)$$

$$z_{\parallel}^{k+1} = z_{\parallel}^k + h\tilde{q}_{\parallel}(t^k, z^k) \quad (36b)$$

$$r^{k+1} = \left( I + h\Gamma(\tilde{\Theta}z_{\perp}^k) \right)^{-1} r^k \quad (36c)$$

where  $h := t^{k+1} - t^k > 0$  is the time-step, and

$$\tilde{q}(t^k, z^k) = \tilde{F}(z^k) + \tilde{\zeta}(t^k, z^k),$$

so that  $\tilde{\zeta}(t^k, z^k) \in \tilde{\mathbf{Z}}(t^k, z^k)$  and  $\tilde{q}(t^k, z^k) = \tilde{q}_{\perp}(t^k, z^k) + \tilde{q}_{\parallel}(t^k, z^k)$ . It thus follows from the maximal monotonicity of the subdifferential map and Minty's theorem [7, Theorem 21.1], that there exists a unique selection  $z_{\perp}^{k+1}$  satisfying (36a). Indeed, the rearrangement of terms in (36a) and the use of (4) (see also [7, Theorem 3.14]), gives an explicit expression for  $z_{\perp}^{k+1}$  satisfying (36a) as

$$z_{\perp}^{k+1} = \text{proj} \left( \tilde{S}(r^{k+1}); z_{\perp}^k + h\tilde{q}_{\perp}(t^k, z^k) \right) = \text{proj} \left( \tilde{\Theta}^{\dagger}S(r^{k+1}); z_{\perp}^k + h\tilde{q}_{\perp}(t^k, z^k) \right), \quad (37)$$

where the second equality is a consequence of Proposition 13-ii) in A. It is worth to remark that the selection (37) is unique and it was possible to compute because of the implicit discretization of the set-value term in (36a). From (37) and in view of Assumption 5 and Proposition 13 it follows that

$$z_{\perp}^k \in \tilde{\Theta}^{\dagger}S(r^k), \text{ and } z^k \in \tilde{S}(r^k) \text{ for all } k \geq 0.$$

It is also worth mentioning that, as  $\Gamma(\tilde{\Theta}z_{\perp}^k)$  is a diagonal matrix with non-negative entries, the inverse in (36c) is well-defined. Moreover, since  $\Gamma(\tilde{\Theta}(z_{\perp}^k))$  is a non-negative diagonal matrix, it follows that all eigenvalues of  $I + h\Gamma(\tilde{\Theta}(z_{\perp}^k))$  lie outside the unitary ball. Recalling that for matrices we consider the induced norm, it is clear that

$$\|(I + h\Gamma(\tilde{\Theta}(z_{\perp}^k)))^{-1}\| \leq 1 \text{ for all } z^k \in \mathbb{R}^n. \quad (38)$$

Hence, the recursion on  $r^k$ , given by (36c), is Schur stable regardless of the value of  $z^k$ .

Let us consider the following family of piecewise linear functions parametrized by  $h$ .

$$z_h(t) = z^k + \frac{t - t^k}{h}(z^{k+1} - z^k) \text{ for } t \in [t^k, t^{k+1}), \quad (39)$$

$$r_h(t) = r^k + \frac{t - t^k}{h}(r^{k+1} - r^k) \text{ for } t \in [t^k, t^{k+1}). \quad (40)$$

It is clear from (39)-(40) that  $z_h$  and  $r_h$  are differentiable almost everywhere in  $[0, T]$ . In what follows it is shown that for  $h$  sufficiently small the sequences  $\{z_h\}_{h>0}$ ,  $\{r_h\}_{h>0}$ ,  $\{\dot{z}_h\}_{h>0}$ ,  $\{\dot{r}_h\}_{h>0}$ , converge in suitable spaces.

### 6.3 Convergence of approximate solutions

Let  $0 = t_0 < t_1 < \dots < t_N = T$ . First, we show that the derivatives  $\dot{z}_h$  and  $\dot{r}_h$  are bounded for  $t \in [0, T]$ . To that end, let  $t \in [0, T] \setminus \{t_0, t_1, \dots, t_N\}$  and consider the following inequality

$$\|\dot{z}_h(t)\| = \frac{1}{h} \|z^{k+1} - z^k\| \leq \frac{1}{h} \|z^{k+1} - z^k - h\tilde{q}(t^k, z^k)\| + \|\tilde{q}(t^k, z^k)\|. \quad (41)$$

Let us focus on the first term on the right-hand side of the inequality. It follows from (36)-(37) and the definition of projection that

$$\begin{aligned} \|z^{k+1} - z^k - h\tilde{q}(t^k, z^k)\| &= \|z_{\perp}^{k+1} + z_{\parallel}^{k+1} - z^k - h\tilde{q}(t^k, z^k)\| \\ &= \|\text{proj}(\tilde{S}(r^{k+1}); z_{\perp}^k + h\tilde{q}_{\perp}(t^k, z^k)) + z_{\parallel}^k + h\tilde{q}_{\parallel}(t^k, z^k) - z^k - h\tilde{q}(t^k, z^k)\| \\ &= \|\text{proj}(\tilde{S}(r^{k+1}); z_{\perp}^k + h\tilde{q}_{\perp}(t^k, z^k)) - z_{\perp}^k - h\tilde{q}_{\perp}(t^k, z^k)\| \\ &= \text{dist}(z_{\perp}^k + h\tilde{q}_{\perp}(t^k, z^k), \tilde{S}(r^{k+1})) \\ &\leq \text{dist}(z_{\perp}^k, \tilde{S}(r^{k+1})) + h\|\tilde{q}_{\perp}(t^k, z^k)\| \\ &= \text{dist}(z_{\perp}^k, \tilde{S}(r^{k+1})) - \text{dist}(z_{\perp}^k, \tilde{S}(r^k)) + h\|\tilde{q}_{\perp}(t^k, z^k)\|, \end{aligned}$$

where we used the fact that  $z_{\perp}^k \in \tilde{S}(r^k)$  on the last line. Making use of Lemma 14 in A yields the upper bound

$$\|z^{k+1} - z^k - h\tilde{q}(t^k, z^k)\| \leq L_S \|r^{k+1} - r^k\| + h\|\tilde{q}_{\perp}(t^k, z^k)\|, \quad (42)$$

and the substitution of (42) into (41) yields the estimate

$$\begin{aligned} \|\dot{z}_h(t)\| &\leq \frac{L_S}{h} \|r^{k+1} - r^k\| + 2\|\tilde{q}(t^k, z^k)\| \\ &\leq L_S \|\dot{r}_h(t)\| + 2(L_{\tilde{F}} + \bar{M}_{\zeta}) \|z^k\| + 2(\|\tilde{F}(0)\| + M_{\zeta}), \end{aligned} \quad (43)$$

where  $L_{\tilde{F}}$  denotes the Lipschitz constant of  $\tilde{F}$  and  $\|\tilde{\zeta}(t^k, z^k)\| \leq \bar{M}_{\zeta} \|z^k\| + M_{\zeta}$  follows from Assumption 1. On the other hand, it follows from (36c) and (40) that

$$\begin{aligned} \|\dot{r}_h(t)\| &= \left\| \Gamma(\tilde{\Theta} z_{\perp}^k) r^{k+1} \right\| \\ &= \left\| \Gamma(\tilde{\Theta} z_{\perp}^k) (I + h\Gamma(\tilde{\Theta} z_{\perp}^k))^{-1} r^k \right\| \\ &\leq \|\Gamma(\tilde{\Theta} z_{\perp}^k)\| \left( \prod_{r=1}^k \left\| (I + h\Gamma(\tilde{\Theta} z_{\perp}^k))^{-1} \right\| \right) \|r(0)\|. \end{aligned}$$

Note that (9) implies that there exist positive constants  $\bar{M}_{\Gamma}, M_{\Gamma}$  such that

$$\|\Gamma(\tilde{\Theta} z_{\perp}^k)\| \leq \bar{M}_{\Gamma} \|z^k\| + M_{\Gamma}.$$

Therefore,

$$\|\dot{r}_h(t)\| \leq \bar{R}(\bar{M}_{\Gamma} \|z^k\| + M_{\Gamma}), \quad (44)$$

where  $\bar{R}$  is given by Assumption 5. Thus, the substitution of (44) into (43), and by setting  $\alpha_1 := L_S \bar{R} \bar{M}_{\Gamma} + 2(\|\tilde{F}(0)\| + M_{\zeta})$ ,  $\alpha_2 := 2L_{\tilde{F}} + 2\bar{M}_{\zeta} + \bar{R} \bar{M}$ , yields

$$\begin{aligned} \|\dot{z}_h(t)\| &\leq \alpha_1 + \alpha_2 \|z^k\| \\ &= \alpha_1 + \alpha_2 \|z_h(t) - (t - t^k) \dot{z}_h(t)\| \\ &\leq \alpha_1 + \alpha_2 (t - t^k) \|\dot{z}_h(t)\| + \alpha_2 \|z_h(t)\| \\ &\leq \alpha_1 + \alpha_2 h \|\dot{z}_h(t)\| + \alpha_2 \|z_h(t)\|. \end{aligned} \quad (45)$$

Let  $0 < \delta < 1$ , then for all  $0 < h < h^* := \frac{1-\delta}{\alpha_2}$  it follows that

$$\begin{aligned}\|\dot{z}_h(t)\| &\leq \frac{\alpha_1}{1-\alpha_2 h} + \frac{\alpha_2}{1-\alpha_2 h} \|z_h(t)\| \\ &\leq \frac{\alpha_1}{\delta} + \frac{\alpha_2}{\delta} \|z_h(t)\|.\end{aligned}\tag{46}$$

The integration on both sides of (46) from  $t = 0$  up to  $t = T$ , yields the estimate

$$\begin{aligned}\|z_h(t) - z_h(0)\| &= \left\| \int_0^t \dot{z}_h(\tau) d\tau \right\| \\ &\leq \int_0^t \|\dot{z}_h(\tau)\| d\tau \leq \frac{\alpha_1}{\delta} t + \frac{\alpha_2}{\delta} \int_0^t \|z_h(\tau)\| d\tau.\end{aligned}\tag{47}$$

Hence,

$$\|z_h(t)\| \leq \frac{\alpha_1}{\delta} t + \|z(0)\| + \frac{\alpha_2}{\delta} \int_0^t \|z_h(\tau)\| d\tau.$$

By Gronwall's inequality, (see *e.g.*, [53, Proposition 3.10]), it follows that

$$\|z_h(t)\| \leq \left( \frac{\alpha_1}{\delta} t + \|z(0)\| \right) e^{\frac{\alpha_2}{\delta} t} \leq \left( \frac{\alpha_1}{\delta} T + \|z(0)\| \right) e^{\frac{\alpha_2}{\delta} T} =: M_T,\tag{48}$$

and therefore,

$$\|\dot{z}_h(t)\| \leq \frac{\alpha_1 + \alpha_2 M_T}{\delta},\tag{49}$$

$$\|\dot{r}_h(t)\| \leq \bar{R}(\bar{M}_\Gamma M_T + M_\Gamma),\tag{50}$$

$$\|\tilde{\zeta}_h(t, z_h(t))\| \leq \bar{M}_\zeta M_T + M_\zeta.\tag{51}$$

Consequently,  $\{z_h\}_{0 < h < h^*}$ ,  $\{\dot{z}_h\}_{0 < h < h^*}$ ,  $\{r_h\}_{0 < h < h^*}$ ,  $\{\dot{r}_h\}_{0 < h < h^*}$ , and  $\{\tilde{\zeta}_h\}_{0 < h < h^*}$  are uniformly bounded for all  $t \in [0, T]$ . In addition, (49) implies that  $\{z_h\}_{0 < h < h^*}$  is equi-Lipschitz in  $[0, T]$ , that is, there is a common Lipschitz constant for all  $z_h$  in the sequence with  $0 < h < h^*$ . Therefore,  $\{z_h\}_{0 < h < h^*}$  is equicontinuous [29, Definition 1.3.17]. Thus, it follows from the Arzelà-Ascoli theorem [10, Theorem 4.25], that there exists a subsequence, still denoted as  $\{z_h\}$ , such that  $z_h \rightarrow z$  uniformly as  $h \downarrow 0$ . Moreover, it follows from (49) that  $\dot{z}_h \in L_1([0, T]; \mathbb{R}^n)$  and it is equi-integrable, see *e.g.*, [29, Definition 3.2.1]. Thus, it follows from the Dunford-Pettis theorem [10, Theorem 4.30], that there exists a subsequence, still denoted as  $\{\dot{z}_h\}$ , such that  $\dot{z}_h \rightharpoonup \dot{z}$  weakly in  $L_1([0, T]; \mathbb{R}^n)$  as  $h \downarrow 0$ . The same arguments apply for the sequences  $\{r_h\}_{0 < h < h^*}$  and  $\{\dot{r}_h\}_{0 < h < h^*}$ . That is,  $r_h \rightarrow r$  uniformly and  $\dot{r}_h \rightharpoonup \dot{r}$  weakly in  $L_1([0, T]; \mathbb{R}^{|\mathcal{E}|m})$  as  $h \downarrow 0$ . Also, as  $\zeta_h \in L_\infty([0, T]; \mathbb{R}^p)$ , it follows from the Banach-Alaoglu-Bourbaki theorem [10, Theorem 3.16], that  $\tilde{\zeta}_h \overset{*}{\rightharpoonup} \tilde{\zeta}$  in the weak\*-topology of  $L_\infty([0, T]; \mathbb{R}^p)$ . Recalling that the uniform limit of sequences of continuous functions is again continuous, and since the sequence of derivatives  $\{\dot{z}_h\}$  and  $\{\dot{r}_h\}$  are equi-integrable, it follows that both limit functions,  $z$  and  $r$ , are absolutely continuous in  $[0, T]$ , [20].

## 6.4 Compliance of the limit functions with the continuous time model

It rests to show that the limit functions of the previous subsection satisfy the differential inclusion (35) for almost all  $t \in [0, T]$ . To that end, let  $\tau_h(t) = t^k$  and  $\tau_h^+(t) = t^{k+1}$  for  $t \in [t^k, t^k + h)$ . Notice that  $\lim_{h \rightarrow 0} \tau_h(t) = \lim_{h \rightarrow 0} \tau_h^+(t) = t$ . Hence, for all  $t \in [0, T]$ ,  $z_\perp^k = [z_h(\tau_h(t))]_\perp$ ,  $r^k = r_h(\tau_h(t))$ , and it follows from Lemma 14 in A that

$$\begin{aligned}\text{dist}(z_\perp(t), \tilde{\Theta}^\dagger S(r(t))) &\leq \text{dist}(z_\perp(t), \tilde{\Theta}^\dagger S(r_h(\tau_h(t)))) + L_S \|r(t) - r_h(\tau_h(t))\| \\ &\leq \|z_\perp(t) - [z_h(\tau_h(t))]_\perp\| + L_S \|r(t) - r_h(\tau_h(t))\|,\end{aligned}$$

Taking the limit as  $h \downarrow 0$  on both sides of the inequality leads us to

$$z_\perp(t) \in \tilde{\Theta}^\dagger S(r(t)), \text{ for all } t \in [0, T],$$

and Proposition 13 implies that

$$z(t) = z_{\perp}(t) + z_{\parallel}(t) \in \tilde{S}(r(t)) \text{ for all } t \in [0, T]. \quad (52)$$

Now it follows from (48)-(49), Assumption 1, and the continuity of  $\tilde{F}$ , that there exists a constant  $0 < \bar{M}_T < \infty$  such that

$$\left\| -\frac{z^{k+1} - z^k}{h} + \tilde{q}(t^k, z^k) \right\| \leq \bar{M}_T, \text{ for all } t^k \in [0, T]. \quad (53)$$

On the other hand, [7, Example 16.49] together with Proposition 13 yield,

$$\begin{aligned} \bar{M}_T \mathbf{N}(\tilde{S}(r^{k+1}); z_{\perp}^{k+1}) \cap \mathbb{B} &= \bar{M}_T \partial \text{dist}(z_{\perp}^{k+1}, \tilde{S}(r^{k+1})) \\ &= \bar{M}_T \partial \text{dist}(z_{\perp}^{k+1}, \tilde{\Theta}^{\dagger} S(r^{k+1})) \end{aligned} \quad (54)$$

Hence, (53) together with (36) imply that

$$-\frac{z_{\perp}^{k+1} - z_{\perp}^k}{h} + \tilde{q}_{\perp}(t^k, z^k) \in \bar{M}_T \partial \text{dist}(z_{\perp}^{k+1}, \tilde{\Theta}^{\dagger} S(r^{k+1})),$$

Therefore,

$$-[\dot{z}_h(\tau_h(t))]_{\perp} + h\tilde{q}_{\perp}(\tau_h(t), z_h(\tau_h(t))) \in \bar{M}_T \partial \text{dist}([z_h(\tau_h^+(t))]_{\perp}, \tilde{\Theta}^{\dagger} S(r_h(\tau_h^+(t)))) ,$$

As  $z_h \rightarrow z$  and  $r_h \rightarrow r$  uniformly as  $h \downarrow 0$ , then also  $[z_h(\tau_h^+(t))]_{\perp} \rightarrow z_{\perp}(t)$  and  $r_h(\tau_h^+(t)) \rightarrow r(t)$  pointwise for all  $t \in [0, T]$  and letting  $l \in L_1([0, T]; \mathbb{R}^n)$  be arbitrary, it follows from Lemma 15 that

$$\begin{aligned} \int_{\Omega} \langle -\dot{z}_{\perp}(t) + \tilde{q}_{\perp}(t, z(t)), l(t) \rangle dt &= \lim_{h \downarrow 0} \int_{\Omega} \langle -[\dot{z}_h(\tau_h(t))]_{\perp} + \tilde{q}_{\perp}(\tau_h(t), z_h(\tau_h(t))), l(t) \rangle dt \\ &\leq \limsup_{h \downarrow 0} \int_{\Omega} \langle -[\dot{z}_h(\tau_h(t))]_{\perp} + \tilde{q}_{\perp}(\tau_h(t), z_h(\tau_h(t))), l(t) \rangle dt \\ &\leq \int_{\Omega} \limsup_{h \downarrow 0} \sigma \left( \bar{M}_T \partial \text{dist}([z_h(\tau_h^+(t))]_{\perp}, \tilde{\Theta}^{\dagger} S(r_h(\tau_h^+(t)))) ; l(t) \right) dt \\ &\leq \int_{\Omega} \sigma \left( \bar{M}_T \partial \text{dist}(z_{\perp}(t); \tilde{\Theta}^{\dagger} S(r(t))) ; l(t) \right) dt. \end{aligned} \quad (55)$$

Since  $z_{\perp}(t) \in \tilde{\Theta}^{\dagger} S(r(t))$  for all  $t \in [0, T]$ , it follows from [13, Proposition III.35] that for almost all  $t \in [0, T]$ ,

$$-[\dot{z}(t)]_{\perp} + \tilde{F}_{\perp}(z(t)) + \tilde{\zeta}_{\perp}(t, z(t)) \in \bar{M}_T \partial \text{dist}(z_{\perp}(t); \tilde{\Theta}^{\dagger} S(r(t))) \subset \mathbf{N}(\tilde{S}(r(t)); z_{\perp}(t)).$$

Analogously, for any  $l \in L_{\infty}([0, T]; \mathbb{R}^n)$

$$\begin{aligned} \int_{\Omega} \langle \tilde{\zeta}(t, z(t)), l(t) \rangle dt &= \lim_{h \downarrow 0} \int_{\Omega} \langle \tilde{\zeta}_h(\tau_h(t), z_h(\tau_h(t))), l(t) \rangle dt \\ &\leq \limsup_{h \downarrow 0} \int_{\Omega} \langle \tilde{\zeta}_h(\tau_h(t), z_h(\tau_h(t))), l(t) \rangle dt \\ &\leq \int_{\Omega} \limsup_{h \downarrow 0} \sigma \left( \tilde{\mathbf{Z}}(\tau_h(t), z_h(\tau_h(t))) ; l(t) \right) dt \\ &\leq \int_{\Omega} \sigma \left( \tilde{\mathbf{Z}}(t, z(t)) ; l(t) \right) dt. \end{aligned}$$

Since  $\mathbf{Z}$  is assumed upper-semicontinuous with compact convex values, it follows from [13, Proposition III.35] that  $\tilde{\zeta}(t, z(t)) \in \tilde{\mathbf{Z}}(t, z(t))$  for almost all  $t \in [0, T]$ . Continuing with the proof, it follows from (36b) and (36c)

that for any  $l(t) \in L_1([0, T]; \mathbb{R}^n)$

$$\begin{aligned} \int_{\Omega} \langle \dot{z}_{\parallel}(t), l(t) \rangle dt &= \lim_{h \rightarrow 0} \int_{\Omega} \langle \tilde{F}_{\parallel}(z_h(\tau_h(t))) + \tilde{\zeta}_{\parallel}(\tau_h(t), z_h(\tau_h(t))), l(t) \rangle dt \\ &= \int_{\Omega} \langle \tilde{F}_{\parallel}(z(t)) + \tilde{\zeta}_{\parallel}(t, z(t)), l(t) \rangle dt, \end{aligned} \quad (56)$$

$$\begin{aligned} \int_{\Omega} \langle \dot{r}(t), l(t) \rangle dt &= \lim_{h \rightarrow 0} \int_{\Omega} \langle -\Gamma(\tilde{\Theta}z_h(\tau_h(t)))r_h(\tau_h^+(t)), l(t) \rangle dt \\ &= \int_{\Omega} \langle -\Gamma(\tilde{\Theta}z_{\perp}(t))r(t), l(t) \rangle dt. \end{aligned} \quad (57)$$

Therefore, for almost all  $t \in [0, T]$ ,

$$\begin{aligned} \dot{z}_{\parallel}(t) &= \tilde{F}_{\parallel}(z(t)) + \tilde{\zeta}_{\parallel}(t, z(t)) \in \tilde{F}_{\parallel}(z(t)) + \tilde{\mathbf{Z}}_{\parallel}(t, z(t)), \\ \dot{r}(t) &= -\Gamma(\tilde{\Theta}z_{\perp}(t))r(t). \end{aligned}$$

This concludes the proof.

## 7 Conclusions and further research

A set-valued coupling protocol achieving robust output convergence in networks of heterogeneous systems is presented. The general case in which each vertex system may have a different dynamical model, with possible different state space, is considered. The proposed scheme uses only the information from the outputs of vertices. It provides a generalization of the funnel coupling strategy in [30] to the multivariable, set-valued setting, and it is also an extension of the adaptable dead-zone coupling in [12]. Two numerical algorithms for the implementation of the coupling law in digital computers are also discussed. The former is based on the implicit discretization of the proposed set-valued coupling leading to a centralized coupling scheme in discrete time. The latter is based on a regularization from the analog circuit and it follows a black-box approach, as no parameter from the nominal dynamics of the vertices is necessary to compute the coupling signal, but it requires a faster sampling frequency in order to maintain the error bounded. Further extensions include the analysis of splitting algorithms to get a distributed way of approximating the coupling (19).

## A Some technical results

In this appendix section some technical results, used mainly in the proof of Theorem 6, are presented.

**Proposition 13.** *For any  $\rho \in \mathbb{R}_+^{|\mathcal{E}|}$ , the set  $\tilde{S}(\rho)$  in (33) is nonempty with nonempty interior whenever  $\rho \in \text{int } \mathbb{R}_+^{|\mathcal{E}|}$ , and*

$$\tilde{S}(\rho) = \tilde{\Theta}^{\dagger} S(\rho) + \text{null } \tilde{\Theta}. \quad (58)$$

Moreover,

$$i) \beta \in \mathbb{R}^n \in \tilde{S}(\rho) \text{ if, and only if, } \beta_{\perp} \in \tilde{\Theta}^{\dagger} S(\rho).$$

$$ii) \mathbf{N}(\tilde{S}(\rho); \beta_{\perp}) = \mathbf{N}(\tilde{\Theta}^{\dagger} S(\rho); \beta_{\perp}) \cap (\text{null } \tilde{\Theta})^{\perp}.$$

*Proof.* First notice that  $\tilde{S}(\rho)$  is nonempty as, by construction, for any  $\rho \in \mathbb{R}_+^{|\mathcal{E}|}$ ,  $0 \in S(\rho)$ , so that  $0 \in \tilde{S}(\rho)$ . Moreover, since  $\tilde{\Theta}$  as a continuous (linear) mapping from  $\mathbb{R}^n$  to  $\mathbb{R}^{|\mathcal{E}|m}$ , then for any  $\rho \in \text{int } \mathbb{R}_+^{|\mathcal{E}|}$ , there exists  $\varepsilon > 0$  such that  $\varepsilon \mathbb{B} \subset \tilde{S}(\rho)$ , and  $\text{int } \tilde{S}(\rho) \neq \emptyset$ . Now, to show that (58) holds, notice that  $\vartheta \in \tilde{S}(\rho)$  if, and only if, there is  $v \in S(\rho)$  such that  $\tilde{\Theta}\vartheta = v$ . Moreover, the last equation holds if, and only if, (see *e.g.*, [8, Proposition 6.1.7]),

$$\vartheta \in \tilde{\Theta}^{\dagger} v + \text{null } \tilde{\Theta},$$

from where (58) easily follows. Now, item  $i$ ) follows from (58), the fact that  $\text{rge } \tilde{\Theta}^\dagger = (\text{null } \tilde{\Theta})^\perp$  [8, Proposition 6.1.6-vii)], and the uniqueness of the splitting  $\beta = \beta_\perp + \beta_\parallel$ . For the second item, let  $\vartheta \in \mathbf{N}(\tilde{S}(\rho); \beta_\perp)$ . Hence, from (58) we have that

$$\begin{aligned} 0 &\geq \langle \vartheta, \eta - \beta_\perp \rangle \text{ for all } \eta \in \tilde{S}(\rho) \\ \Leftrightarrow 0 &\geq \langle \vartheta, \eta_\parallel + \eta_\perp - \beta_\perp \rangle \text{ for all } \eta_\parallel \in \text{null } \tilde{\Theta}, \eta_\perp \in \tilde{\Theta}^\dagger S(\rho). \end{aligned} \quad (59)$$

Since (32) implies that  $\vartheta \in \text{rge}(\tilde{\Theta}^\dagger) = (\text{null } \tilde{\Theta})^\perp$ , it follows that  $\langle \vartheta, \eta_\parallel \rangle = 0$  for all  $\eta_\parallel \in \text{null } \tilde{\Theta}$  and then (59) implies that  $\mathbf{N}(\tilde{S}(\rho); \beta_\perp) \subseteq \mathbf{N}(\tilde{\Theta}^\dagger S(\rho); \beta_\perp) \cap (\text{null } \tilde{\Theta})^\perp$ . For the converse, take  $\vartheta \in \mathbf{N}(\tilde{\Theta}^\dagger S(\rho); \beta_\perp) \cap (\text{null } \tilde{\Theta})^\perp$ . Then,  $\langle \vartheta, \eta_\parallel + \eta_\perp - \beta_\perp \rangle = \langle \vartheta, \eta_\perp - \beta_\perp \rangle$  for any  $\eta_\parallel \in \text{null } \tilde{\Theta}$ , and the equivalence (59) implies that  $\mathbf{N}(\tilde{S}(\rho); \beta_\perp) \supseteq \mathbf{N}(\tilde{\Theta}^\dagger S(\rho); \beta_\perp) \cap (\text{null } \tilde{\Theta})^\perp$ . This concludes the proof.  $\square$

Let  $\mathcal{U}, \mathcal{V} \subset \mathbb{R}^l$  be two nonempty, compact sets. The Pompeiu-Hausdorff distance between  $\mathcal{U}$  and  $\mathcal{V}$ , denoted as  $d_H(\mathcal{U}, \mathcal{V})$ , is given as

$$d_H(\mathcal{U}, \mathcal{V}) = \sup_{x \in \mathcal{U} \cup \mathcal{V}} |\text{dist}(x, \mathcal{U}) - \text{dist}(x, \mathcal{V})|, \quad (60)$$

or equivalently

$$d_H(\mathcal{U}, \mathcal{V}) = \inf \{ \varepsilon \geq 0 \mid \mathcal{U} \subseteq \mathcal{V} + \varepsilon \mathbb{B}_l, \text{ and } \mathcal{V} \subseteq \mathcal{U} + \varepsilon \mathbb{B}_l \}, \quad (61)$$

Thus, it follows from (60), that for any  $\eta \in \mathbb{R}^l$ ,

$$|\text{dist}(\eta, \mathcal{U}) - \text{dist}(\eta, \mathcal{V})| \leq d_H(\mathcal{U}, \mathcal{V}). \quad (62)$$

The Pompeiu-Hausdorff distance measures how far two sets are from each other and it provides a notion of continuity to the motion of time-varying sets. Roughly speaking, the following lemma shows that the set  $\tilde{S}$  in (33) ‘‘moves in a Lipschitz continuous way’’.

**Lemma 14.** *For the set  $\tilde{S}$  defined in (33) there exists a positive constant  $L_S > 0$  such that for any two  $\eta_1, \eta_2 \in \mathbb{R}^{|E|}$  and any  $\beta_\perp \in (\text{null } \tilde{\Theta})^\perp$ , the following holds*

$$|\text{dist}(\beta_\perp, \tilde{S}(\eta_1)) - \text{dist}(\beta_\perp, \tilde{S}(\eta_2))| \leq L_S \|\eta_1 - \eta_2\|. \quad (63)$$

*Proof.* It follows from Proposition 13 and (62) that,

$$\begin{aligned} |\text{dist}(\beta_\perp, \tilde{S}(\eta_1)) - \text{dist}(\beta_\perp, \tilde{S}(\eta_2))| &= |\text{dist}(\beta_\perp, \tilde{\Theta}^\dagger S(\eta_1)) - \text{dist}(\beta_\perp, \tilde{\Theta}^\dagger S(\eta_2))| \\ &\leq d_H(\tilde{\Theta}^\dagger S(\eta_1), \tilde{\Theta}^\dagger S(\eta_2)). \end{aligned} \quad (64)$$

Additionally, (14) implies that  $\tilde{\Theta}^\dagger S(\eta_2) \subseteq \tilde{\Theta}^\dagger S(\eta_1) + \|\tilde{\Theta}^\dagger(\eta_1 - \eta_2)\|_\infty \mathbb{B}$  and  $\tilde{\Theta}^\dagger S(\eta_1) \subseteq \tilde{\Theta}^\dagger S(\eta_2) + \|\tilde{\Theta}^\dagger(\eta_1 - \eta_2)\|_\infty \mathbb{B}$ . Consequently,

$$d_H(\tilde{\Theta}^\dagger S(\eta_1), \tilde{\Theta}^\dagger S(\eta_2)) \leq \|\tilde{\Theta}^\dagger(\eta_1 - \eta_2)\|_\infty \leq \|\tilde{\Theta}^\dagger(\eta_1 - \eta_2)\|_2 \leq \|\tilde{\Theta}^\dagger\| \|\eta_1 - \eta_2\|. \quad (65)$$

Finally, the combination of (64) and (65) leads us to (63). The proof is thus complete.  $\square$

**Lemma 15.** *Let  $\{\beta_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$ ,  $\{\rho_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^{|E|}$  be two convergent sequences such that  $\beta_k \rightarrow \beta$  and  $\rho_k \rightarrow \rho$  as  $k \rightarrow \infty$ . Then, for any  $q \in \mathbb{R}^n$ ,*

$$\limsup_{k \rightarrow \infty} \sigma \left( \partial \text{dist}(\beta_k, \tilde{\Theta}^\dagger S(\rho_k)); q \right) \leq \sigma \left( \partial \text{dist}(\beta, \tilde{\Theta}^\dagger S(\rho)); q \right). \quad (66)$$

*Proof.* Let  $p \in \partial \text{dist}(\beta_k, \tilde{\Theta}^\dagger S(\rho_k))$ , it thus follows from Lemma 14 that, for any  $s \in \mathbb{R}^n$ ,

$$\begin{aligned} \langle p, s - \beta_k \rangle &\leq \text{dist}(s, \tilde{\Theta}^\dagger S(\rho_k)) - \text{dist}(\beta_k, \tilde{\Theta}^\dagger S(\rho_k)) \\ &\leq \text{dist}(s, \tilde{\Theta}^\dagger S(\rho)) - \text{dist}(\beta_k, \tilde{\Theta}^\dagger S(\rho)) + 2L_S \|\rho_k - \rho\|. \end{aligned}$$

Hence, for any  $\varepsilon > 0$  there exists  $k^* \in \mathbb{N}$  such that, for all  $k \geq k^*$ ,

$$\partial \text{dist}(\beta_k, \tilde{\Theta}^\dagger S(\rho_k)) \subset \partial \text{dist}(\beta_k, \tilde{\Theta}^\dagger S(\rho)) + \varepsilon \mathbb{B}. \quad (67)$$

For any nonempty, closed, convex set  $\tilde{\Theta}^\dagger S(\rho) \subset \mathbb{R}^n$  the distance function  $\text{dist}(\cdot, \tilde{\Theta}^\dagger S(\rho))$  is a proper, lower semicontinuous, convex function [7, Corollary 12.12 and Proposition 12.14]. Hence,  $\partial \text{dist}(\cdot, \tilde{\Theta}^\dagger S(\rho))$  is an upper-semicontinuous set-valued map. It thus follows from (67) that, for any  $\varepsilon > 0$  there is  $k^* \in \mathbb{N}$  such that for all  $k \geq k^*$ ,

$$\partial \text{dist}(\beta_k, \tilde{\Theta}^\dagger S(\rho_k)) \subset \partial \text{dist}(\beta, \tilde{\Theta}^\dagger S(\rho_k)) + \varepsilon \mathbb{B}. \quad (68)$$

Thus, for any  $\varepsilon > 0$ , the support function satisfies,

$$\begin{aligned} \sigma(\partial \text{dist}(\beta_k, \tilde{\Theta}^\dagger S(\rho_k)); q) &= \sup_{p \in \partial \text{dist}(\beta_k, \tilde{\Theta}^\dagger S(\rho_k))} \langle p, q \rangle \\ &\leq \sup_{p \in \partial \text{dist}(\beta, \tilde{\Theta}^\dagger S(\rho))} \langle p, q \rangle + \varepsilon \sup_{b \in \mathbb{B}} \langle b, q \rangle \\ &= \sigma(\partial \text{dist}(\beta, \tilde{\Theta}^\dagger S(\rho)); q) + \varepsilon \|q\|. \end{aligned} \quad (69)$$

Finally, taking the lim sup on both sides of (69) leads us to the desired result.  $\square$

## References

- [1] V. Acary and B. Brogliato. *Numerical Methods for Nonsmooth Dynamical Systems*. Springer-Verlag, Berlin, 2008.
- [2] V. Acary and B. Brogliato. Implicit Euler numerical scheme and chattering-free implementation of sliding mode systems. *Systems & Control Letters*, 59(5):284–293, 2010.
- [3] H. Ahmed, R. Ushirobira, D. Efimov, L. Fridman, and Y. Wang. Output global oscillatory synchronization of heterogeneous systems. *International Journal of Control*, 94(7):1982–1993, 2021.
- [4] Z. Aminzare and E. D. Sontag. Synchronization of diffusively-connected nonlinear systems: results based on contractions with respect to general norms. *IEEE Transactions on Network Science and Engineering*, 1(2):91–106, 2014.
- [5] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou. Synchronization in complex networks. *Physics Reports*, 469:93–153, 2008.
- [6] J. Bastien. Convergence order of implicit Euler numerical scheme for maximal monotone differential inclusions. *Zeitschrift für angewandte Mathematik und Physik*, 64(4):955–966, 2013.
- [7] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer New York, 2011.
- [8] D. S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas*. Princeton University Press, New Jersey, USA, 2009.
- [9] A. Bizyaeva, A. Franci, and N. E. Leonard. Nonlinear opinion dynamics with tunable sensitivity. *IEEE Transactions on Automatic Control*, 68(3):1415–1430, 2023.
- [10] H. Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York, USA, 2010.



- [11] C. I. Byrnes, A. Isidori, and J. C. Willems. Passivity, feedback equivalence, and the global stabilization of minimum phase nonlinear systems. *IEEE Transaction on Automatic Control*, 36(11):1228–1240, 1991.
- [12] G. Casadei, D. Astolfi, A. Alessandri, and L. Zaccarian. Synchronization in networks of identical nonlinear systems via dynamic dead zones. *IEEE Control Systems Letters*, 3(3):667–672, 2019.
- [13] C. Castaing and M. Valadier. *Convex Analysis and Measurable Multifunctions*. Springer-Verlag, Germany, 1977.
- [14] L. Chen, C. Qiu, and H. B. Huang. Synchronization with on-off coupling: role of time scales in network dynamics. *Physical Review E*, 79(045101), 2009.
- [15] M. Coraggio, P. DeLellis, and M. di Bernardo. Convergence and synchronization in networks of piecewise-smooth systems via distributed discontinuous coupling. *Automatica*, 129:1–9, 2021.
- [16] M. Coraggio, P. DeLellis, and M. di Bernardo. Distributed discontinuous coupling for convergence in heterogeneous networks. *IEEE Control Systems Letters*, 5(3):1037–1042, 2021.
- [17] P. DeLellis, M. di Bernardo, T. E. Goroehowski, and G. Russo. Synchronization and control of complex networks via contraction, adaptation and evolution. *IEEE Circuits and Systems Magazine*, pages 64–82, 2010.
- [18] P. DeLellis, M. di Bernardo, and G. Russo. On QUAD, Lipschitz, and contracting vector fields for consensus and synchronization of networks. *IEEE Transactions on Circuits and Systems I*, 58(3):576–583, 2011.
- [19] F. Dörfler and F. Bullo. Synchronization in complex networks of phase oscillators: a survey. *Automatica*, 50:1539–1564, 2014.
- [20] P. M. Fitzpatrick and B. R. Hunt. Absolute continuity of a functions and uniform integrability of its divided differences. *The American Mathematical Monthly*, 122(4):362–366, 2015.
- [21] L. V. Gambuzza, M. Frasca, and V. Latora. Distributed control of synchronization of a group of network nodes. *IEEE Transactions on Automatic Control*, 64(1):365–372, 2019.
- [22] D. Ghosh, M. Frasca, A. Rizzo, S. Majhi, S. Rakshit, K. Alfaro-Bittner, and S. Boccaletti. The synchronized dynamics of time-varying networks. *Physics Reports*, 949:1–63, 2022.
- [23] C. Godsil and G. Royle. *Algebraic Graph Theory*. Springer, New York, USA, 2001.
- [24] Z.-H. Guan, Z.-W. Liu, G. Feng, and Y.-W. Wang. Synchronization of complex dynamical networks with time-varying delays via impulsive distributed control. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 57:2182–2195, 2010.
- [25] T. Haddad. Differential inclusions governed by a state dependent sweeping process. *International Journal of Difference Equations*, 8(1):63–70, 2013.
- [26] N. Jiang, X. Liu, W. Yu, and J. Shen. Finite-time stochastic synchronization of genetic regulatory networks. *Neurocomputing*, 167:314–321, 2015.
- [27] C. Jin, Z. Wang, L. Gong, M. Xiao, and G.-P. Jiang. Quasi-synchronization of heterogeneous Lur’e networks with uncertain parameters and impulsive effect. *Neurocomputing*, 482:252–263, 2022.
- [28] J. Kim, J. Yang, H. Shim, J.-S. Kim, and Seo J. H. Robustness of synchronization of heterogeneous agents by strong coupling and a large number of agents. *IEEE Transactions on Automatic Control*, 61(10):3096–3102, 2016.
- [29] A. J. Kurdila and M. Zabaranin. *Convex Functional Analysis*. Systems & Control: Foundations & Applications. Birkhäuser Verlag, Germany, 2005.

- [30] J. G. Lee, S. Trenn, and H. Shim. Synchronization with prescribed transient behavior: heterogeneous multi-agent systems under funnel coupling. *Automatica*, 141:110276, 2022.
- [31] A. Levant. Higher-order sliding modes, differentiation and output-feedback control. *International Journal of Control*, 76(9-10):924–941, 2003.
- [32] M. Lodi, F. Della Rossa, F. Sorrentino, and M. Storace. Analyzing synchronized clusters in neuron networks. *Scientific Reports*, 10(16336), 2020.
- [33] F. A. Miranda-Villatoro and F. Castaños. Robust output regulation of strongly linear passive systems with multivalued maximally monotone controls. *IEEE Transactions on Automatic Control*, 62(1):238–249, 2017.
- [34] J. M. Montenbruck, M. Bürger, and F. Allgöwer. Practical synchronization with diffusive couplings. *Automatica*, 53:235–243, 2015.
- [35] J.-J. Moreau. Evolution problem associated with a moving convex set in a Hilbert space. *Journal of differential equations*, 26(3):347–374, 1977.
- [36] A. E. Motter, S. A. Myers, M. Anghel, and T. Nishikawa. Spontaneous synchrony in power-grid networks. *Nature Physics*, 9(3):191–197, 2013.
- [37] C. Murguia, R. H. B. Fey, and H. Nijmeijer. Network synchronization by dynamic diffusive coupling. *International Journal of Bifurcation and Chaos*, 23(4), 2013.
- [38] M. E. J. Newman and D. J. Watts. Renormalization group analysis of the small-world network model. *Physics Letters A*, 263:341–346, 1999.
- [39] E. Panteley and A. Loría. Synchronization and dynamic consensus of heterogeneous networked systems. *IEEE Transactions on Automatic Control*, 62(8):3758–3773, 2017.
- [40] N. Papageorgiou. Differential inclusions with state constraints. *Proceedings of the Edinburgh Mathematical Society*, 32:81–98, 1989.
- [41] L. M. Pecora and T. L. Carroll. Master stability functions for synchronized coupled systems. *Physical Review Letters*, 8(10):2019–2112, 1998.
- [42] S. Phillips and R. G. Sanfelice. Robust distributed synchronization of networked linear systems with intermittent information. *Automatica*, 105:323–333, 2019.
- [43] A. Pikovsky, M. Rosenblum, and J. Kurths. *Synchronization: A universal concept in nonlinear sciences*. Cambridge University Press, 2001.
- [44] A. Y. Pogromsky. Passivity based design of synchronizing systems. *International Journal of Bifurcation and Chaos*, 8(2):295–319, 1998.
- [45] A. V. Proskurnikov, F. Zhang, M. Cao, and J. M. A. Scherpen. A general criterion for synchronization of incrementally dissipative nonlinear coupling agents. In *European Control Conference (ECC)*, pages 581–586, Linz, Austria, 2015.
- [46] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer-Verlag, 2009.
- [47] G. Russo and J.-J. E. Slotine. Symmetries, stability, and control in nonlinear systems and networks. *Physical Review E*, 84(041929), 2011.
- [48] Y. Shtessel, C. Edwards, L. Fridman, and A. Levant. *Sliding Mode Control and Observation*. Birkhäuser, 2014.
- [49] D. J. Stilwell, E. M. Bollt, and D. G. Roberson. Sufficient conditions for fast switching synchronization in time-varying network topologies. *SIAM Journal on Applied Dynamical Systems*, 5(1):140–156, 2006.

- [50] H. L. Trentelman, K. Takaba, and N. Monshizadeh. Robust synchronization of uncertain linear multi-agent systems. *International Journal of Robust and Nonlinear Control*, 58(6):1511–1523, 2013.
- [51] V. I. Utkin, J. Guldner, and J. Shi. *Sliding Mode Control in Electro-Mechanical Systems*, volume 34. CRC press, 2009.
- [52] W. Wang and J.-J. E. Slotine. On partial contraction analysis for coupled nonlinear oscillators. *Biological Cybernetics*, 92:2004, 2004.
- [53] E. Zeidler. *Nonlinear Functional Analysis and its Applications I: Fixed-Point Theorems*. Springer-Verlag, New York, USA, 1986.
- [54] J. Zhao, D. J. Hill, and T. Liu. Synchronization of complex dynamical networks with switching topology: A switched system point of view. *Automatica*, 45:2502–2511, 2009.
- [55] J. Zhao, D. J. Hill, and T. Liu. Global bounded synchronization of general dynamical networks with nonidentical nodes. *IEEE Transactions on Automatic Control*, 57(10):2656–2662, 2012.
- [56] J. Zhou, Y. Zou, S. Guan, Z. Liu, and S. Boccaletti. Synchronization in slowly switching networks of coupled oscillators. *Nature Scientific Reports*, 6(35979), 2016.
- [57] Y. Zhu, W. Ren, W. Yu, and G. Wen. Distributed resource allocation over directed graphs via continuous-time algorithms. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(2):1097–1106, 2015.