



**HAL**  
open science

# Consensus Formation in First-Order Graphon Models with Time-Varying Topologies

Benoît Bonnet, Nastassia Pouradier Duteil, Mario Sigalotti

► **To cite this version:**

Benoît Bonnet, Nastassia Pouradier Duteil, Mario Sigalotti. Consensus Formation in First-Order Graphon Models with Time-Varying Topologies. *Mathematical Models and Methods in Applied Sciences*, 2022, 32 (11), pp.2121-2188. hal-03418399v2

**HAL Id: hal-03418399**

**<https://hal.science/hal-03418399v2>**

Submitted on 16 Oct 2022

**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.

# Consensus Formation in First-Order Graphon Models with Time-Varying Topologies

Benoît Bonnet\*, Nastassia Pouradier Duteil† and Mario Sigalotti‡

October 16, 2022

## Abstract

In this article, we investigate the asymptotic formation of consensus for several classes of time-dependent cooperative graphon dynamics. After motivating the use of this type of macroscopic models to describe multi-agent systems, we adapt the classical notion of scrambling coefficient to this setting, leverage it to establish sufficient conditions ensuring the exponential convergence to consensus with respect to the  $L^\infty$ -norm topology. We then shift our attention to consensus formation expressed in terms of the  $L^2$ -norm, and prove three different consensus result for symmetric, balanced and strongly connected topologies, which involve a suitable generalisation of the notion of algebraic connectivity to this infinite-dimensional framework. We then show that, just as in the finite-dimensional setting, the notion of algebraic connectivity that we propose encodes information about the connectivity properties of the underlying interaction topology. We finally use the corresponding results to shed some light on the relation between  $L^2$ - and  $L^\infty$ -consensus formation, and illustrate our contributions by a series of numerical simulations.

**Keywords :** Multi-Agent Systems, Graphon Dynamics, Consensus Formation, Algebraic Connectivity, Scrambling Coefficient, Persistence Conditions

**MSC2020 Subject Classification :** 05C63, 05C90, 37L15, 93A16

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Preliminaries</b>	<b>6</b>
2.1	Measure theory and integration . . . . .	6
2.2	Graph and spectral theory . . . . .	7
2.3	Multi-agent systems and graphon models . . . . .	9
<b>3</b>	<b>Consensus formation via diameter estimates</b>	<b>13</b>
3.1	Generalised scrambling coefficient for graphon models . . . . .	14
3.2	Exponential consensus formation for topologies with persistent scramblings . . . . .	16
<b>4</b>	<b>Consensus formation via energy estimates</b>	<b>20</b>
4.1	Generalised algebraic connectivity for graphon models . . . . .	20
4.2	Exponential consensus formation for algebraically persistent topologies . . . . .	25
4.2.1	Symmetric nonlinear interaction topologies . . . . .	26
4.2.2	Balanced linear interactions topologies . . . . .	30
4.2.3	Disjoint unions of strongly connected linear topologies . . . . .	32

---

\*Inria Paris & Laboratoire Jacques-Louis Lions, Sorbonne Université, Université Paris-Diderot SPC, CNRS, Inria, 75005 Paris, France *Email:* benoit.a.bonnet@inria.fr

†Inria Paris & Laboratoire Jacques-Louis Lions, Sorbonne Université, Université Paris-Diderot SPC, CNRS, Inria, 75005 Paris, France *Email:* nastassia.pouradierduteil@inria.fr

‡Inria Paris & Laboratoire Jacques-Louis Lions, Sorbonne Université, Université Paris-Diderot SPC, CNRS, Inria, 75005 Paris, France *Email:* mario.sigalotti@inria.fr

<b>5 Further discussions around connectivity and consensus formation</b>	<b>34</b>
5.1 Link between algebraic and graphon connectivity . . . . .	34
5.2 Equivalence between $L^2$ - and $L^\infty$ -consensus for strongly connected graphons . . . . .	38
5.3 Numerical illustrations . . . . .	39

## 1 Introduction

Since its appearance at the end of the nineteen nineties, the analysis of self-organisation and pattern formation in large groups of interacting individuals has become a prominent topic in applied mathematics. This subject was initially brought forth by the pioneering articles [36, 75] on statistical physics oriented models for biological systems, and subsequently cemented by a wealth of contributions in the fields of automation theory and engineering, see e.g. [6, 12, 13, 45, 52, 57, 59, 74] and references therein. In the midst of this broad academical trend, a research current led by the works of Hegselman and Krause [41, 46] on bounded confidence models, and the groundbreaking papers of Cucker and Smale [27, 28] on emergent behaviours, started to focus more specifically on the problem of *consensus formation*. In the meantime, many of the ideas that came into existence around these questions found their way to a large score of application areas, such as herds and flocks analysis [5, 7], aggregation models in biology [25], decentralised control of fleets of autonomous vehicles [6, 17] or pedestrian dynamics [26, 66]. Besides consensus problems, which by far attracted the most attention in the literature, other families of relevant and perhaps more intricate stable patterns – such as mills [11, 23] or rings [9] – are being investigated by communities centered around multi-agent dynamics.

Motivated by the apparent ubiquity of clustering behaviours in applied sciences, a growing literature has been devoted to the precise mathematical understanding of the mechanisms subtending this type of pattern formation in the class of *multi-agent systems*. These latter colloquially refer to systems of  $N \geq 1$  agents represented by points in a given configuration space – in our context the euclidean space  $\mathbb{R}^d$  –, which evolve according to a coupled dynamics of the form

$$\dot{x}_i(t) = \frac{1}{N} \sum_{j=1}^N \psi_{ij}(t, x_i(t) - x_j(t)). \quad (1.1)$$

Here, the vector  $(x_1(t), \dots, x_N(t)) \in (\mathbb{R}^d)^N$  represents the collection of all the states of the agents at some time  $t \geq 0$ , while the maps  $\psi_{ij} : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  defined for each pair of indices  $i, j \in \{1, \dots, N\}$ , encode nonrepulsive pairwise interactions between agents, which may depend on time, on their relative distance and orientation, as well as on their respective labels. In this context, one is usually interested in understanding under which circumstances solutions of (1.1) *converge to consensus*, namely, when there does exist an element  $x^\infty \in \mathbb{R}^d$  such that

$$\lim_{t \rightarrow +\infty} |x_i(t) - x^\infty| = 0,$$

for each agent label  $i \in \{1, \dots, N\}$ . This question has been the object of a tremendously large amount of work – see e.g. [13, 19, 20, 29, 44, 45, 52, 59, 60, 64, 72, 73] and references therein –, whose number can be explained by the necessity to tackle a wide variety of specific interaction rules described by the maps  $(\psi_{ij})_{1 \leq i, j \leq N}$ . Some of the farthest-reaching results on this topic can be attributed to Moreau [59] as well as to Olfati-Saber and Murray [64], while the main approaches and proof strategies to study asymptotic clustering are outlined in the reference survey [61] by Motsch and Tadmor.

Even though the literature devoted to consensus formation is very broad, quantitative convergence results – typically obtained by means of Lyapunov arguments that do not rely on LaSalle’s invariance principle –, are almost exclusively based on *diameter* or *energy* decay estimates. In both cases, the convergence towards consensus is dictated by a scalar quantity, which carries information about the connectivity properties of the system. In the first scenario, diameter contraction estimates usually depend on the so-called *scrambling coefficient* of the topology. The latter is positive whenever each pair of agents in the system is either directly interacting, or following a common third party agent, and thus encodes information about “local” communications in the system. It first appeared in the theory of stochastic matrices as a way of quantifying ergodicity properties of Markov chains, and is

discussed in the reference monograph of Seneta [71]. In the second case, energy dissipation estimates are often expressed in terms of *algebraic connectivity*. This quantity – which is widespread in algebraic graph theory – corresponds to the second smallest eigenvalue of the so-called *graph-Laplacian* matrix associated with the interaction topology. It is known to provide information about the number of connected components in the interaction graph, and allows to quantify the minimum communication strength between agents over all connected components. As such, the algebraic connectivity carries a fairly “global” information about the structure of the topology, and a detailed account on its many properties in the context of undirected graph can be found in the reference article of Mohar [58].

In addition to identifying conditions ensuring the formation of consensus for systems of the form (1.1), it is also very natural to try to understand the behaviour of such patterns in large interacting systems, i.e. when the number  $N \geq 1$  of agents goes to infinity. This problem is all the more relevant since in the common application areas of cooperative systems, the number of agents can range between tens of thousands and millions, making frontal resolution potentially arduous from a numerical standpoint. In the situation where the interaction maps  $(\psi_{ij})_{1 \leq i, j \leq N}$  do not depend on the agent labels  $i, j \in \{1, \dots, N\}$  and are all equal to a common function  $\psi : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ , it was shown in [22, 38, 39] that meaningful kinetic approximations formulated in terms of *mean-field limits* could be proposed for (1.1), see also [67]. These latter are formulated as transport equations of the form

$$\partial_t \mu(t) + \operatorname{div}_x(\Psi(t, \mu(t))\mu(t)) = 0, \quad (1.2)$$

where the nonlocal interaction velocities driving the system are given by

$$\Psi(t, \mu, x) := \int_{\mathbb{R}^d} \psi(t, y - x) d\mu(y).$$

In this context, a rigorous embedding can be established between the microscopic system (1.1) and its macroscopic approximation (1.2), by considering the *empirical measures* defined by

$$\mu_N(t) := \frac{1}{N} \sum_{i=1}^N \delta_{x_i(t)}$$

which are supported on the discrete positions  $(x_1(t), \dots, x_N(t)) \in (\mathbb{R}^d)^N$  of the agents at time  $t \geq 0$ . These infinite-dimensional models are usually studied by means of a variety of tools coming from optimal transport theory – see for instance the survey [24] –, and we point the reader to [15, 65] for standard Cauchy-Lipschitz well-posedness results for nonlocal dynamics of the form (1.2).

However when the interaction kernels  $(\psi_{ij})_{1 \leq i, j \leq N}$  depend explicitly on the agent labels, the standard mean-field approach is not applicable any more. Indeed, the latter can only account for systems whose dynamics are invariant under permutations, which greatly limits the scope of admissible interaction models. Motivated by this observation, several recent contributions have been aiming at studying macroscopic approximations of (1.1) which allow for more general and possibly asymmetric interaction functions. The corresponding approaches are based on the theory of *graph limits* introduced by Lovász and Segedy in [50] – see also the corresponding monograph [49] –, and popularised by Medvedev in [55, 56]. In this context, the discrete set of labels  $\{1, \dots, N\}$  allowing to discriminate between agents becomes a continuous unit interval  $I := [0, 1]$ , and the state of the system at time  $t \geq 0$  is represented by a Lebesgue measurable function  $x(t) \in L^2(I, \mathbb{R}^d)$  which maps to each generalised index  $i \in I$  a position  $x(t, i) \in \mathbb{R}^d$ . In this context, the *graphon dynamics* corresponding to (1.1) writes as

$$\partial_t x(t, i) = \int_I \psi(t, i, j, x(t, i) - x(t, j)) dj, \quad (1.3)$$

where  $\psi : \mathbb{R}_+ \times I \times I \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a general function encoding interactions between agents, depending on time, on their relative position and on their respective labels. At present, graphon models of the form (1.3) are starting to emerge in the literatures related to multi-agent systems [4, 16], control theory [10, 37], mean-field games [18, 21], and data sciences [70, 76]. We also point the reader to the papers [16, 47], which to the best of our knowledge are currently the only ones investigating consensus

formation problems for graphon dynamics. We also mention the much older work [12], in which the authors intuited a consensus model with a continuum of agent labels without relying on, or referring to, the terminology of dense graphs.

In this article, our goal is to investigate several families of sufficient conditions ensuring the formation of consensus for solutions of generalised Cucker-Smale graphon models of the form

$$\partial_t x(t, i) = \int_I a(t, i, j) \phi(|x(t, i) - x(t, j)|) (x(t, j) - x(t, i)) dj. \quad (1.4)$$

Here, the map  $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is a smooth and bounded nonlinear interaction kernel accounting for the *distance-based* magnitude of communications between agents, while the time-dependent kernel  $a : \mathbb{R}_+ \times I \times I \rightarrow [0, 1]$  encodes the *interaction topology* of the system. We stress that the latter is not symmetric a priori, so that the fine structure of the interaction topology will be the main factor determining what kind of convergence result one may hope to recover. Taking inspiration from the classical dichotomy between  $L^\infty$ - and  $L^2$ -dissipation estimates in the consensus literature, we start in Section 3 by studying sufficient conditions ensuring the contraction of the diameter of the system

$$\mathcal{D}(t) := \operatorname{ess\,sup}_{i, j \in I} |x(t, i) - x(t, j)|,$$

defined for all times  $t \geq 0$  along solutions of (1.4). To this end, we introduce in Section 3.1 the following macroscopic generalisation for graphon models of the classical scrambling coefficient

$$\eta(\mathcal{A}(t)) := \operatorname{ess\,inf}_{i, j \in I} \int_I \min \{a(t, i, k), a(t, j, k)\} dk.$$

Here,  $\mathcal{A}(t) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  denotes the kernel-type *adjacency operator* generated by  $a(t) \in L^\infty(I \times I, [0, 1])$ , and can be seen as the counterpart of the notion of adjacency matrix for standard graphs. By following an original methodology relying on the theorems of Danskin and Scorza-Dragoni, combined with a key geometric lemma on the maximisers of relaxed diameters, we show in Theorem 3.3 that

$$\mathcal{D}(t) \leq \mathcal{D}(0) \exp \left( -\gamma_R \int_0^t \eta(\mathcal{A}(s)) ds \right), \quad (1.5)$$

for all times  $t \geq 0$ , where  $\gamma_R > 0$  is a constant related to the minimum value taken by the nonlinear kernel  $\phi(\cdot)$  when the  $L^\infty$ -norm of the initial datum of (1.4) is bounded by some constant  $R > 0$ . In addition to its novelty in the context of graphon models, this contraction estimate also generalises those already available for discrete multi-agent systems, as it does not require that the adjacency matrix be stochastic, see e.g. [60, 61]. Building on (1.5), we proceed by showing in Theorem 3.6 that solutions of (1.4) exponentially converge to consensus in the  $L^\infty$ -norm topology as soon as the scrambling coefficients  $\eta(\mathcal{A}(\cdot))$  are *persistent*, in the sense that

$$\frac{1}{\tau} \int_t^{t+\tau} \eta(\mathcal{A}(s)) ds \geq \mu, \quad (1.6)$$

for all times  $t \geq 0$  and some fixed parameters  $(\tau, \mu) \in \mathbb{R}_+^* \times (0, 1]$ . This type of integral condition is fairly standard in stability theory for time-varying systems, as evidenced by the monographs [53, 63].

In Section 4, we shift our focus to energy dissipation estimates involving the standard deviations

$$X(t) := \left( \frac{1}{2} \int_I \int_I |x(t, i) - x(t, j)|^2 dj di \right)^{1/2},$$

defined for all times  $t \geq 0$ , and which allow for a convenient investigation of  $L^2$ -consensus formation. To this end, we start in Section 4.1 by reformulating the dynamics (1.4) as the differential equation

$$\dot{x}(t) = -\mathbb{L}(t, x(t))x(t), \quad (1.7)$$

posed in the Hilbert space  $L^2(I, \mathbb{R}^d)$ . Therein, we introduced the so-called *graph-Laplacian operators*  $\mathbb{L}(t, x) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  associated with the nonlinear interaction topology of the system. Taking

inspiration from the work of Wu [79, 80, 81] on the adaptation of the notion of *algebraic connectivity* to directed graphs, we propose several generalisations of this object to the setting of graphon dynamics. For *symmetric* and *balanced* interaction topologies – see Definition 4.4 below –, we define the algebraic connectivity of an interaction topology whose graph-Laplacian operator is  $\mathbb{L} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  as

$$\lambda_2(\mathbb{L}) := \inf_{x \in \mathcal{C}^\perp} \frac{\langle \mathbb{L}x, x \rangle_{L^2(I, \mathbb{R}^d)}}{\|x\|_{L^2(I, \mathbb{R}^d)}^2}, \quad (1.8)$$

where  $\mathcal{C} \subset L^2(I, \mathbb{R}^d)$  is the so-called *consensus manifold*. In the subtler and more technical case in which the interaction topology is a disjoint union of strongly connected components – understood in the sense of Definition 4.5 below –, it is again possible to define the algebraic connectivity  $\lambda_2(\mathbb{L})$  in a fashion that is similar to (1.8) by renormalising  $\mathbb{L} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  using a suitable function belonging to the kernel of its adjoint. Let it be noted that the existence of this canonical weight for graphon models is a consequence of the recent work [16].

We then proceed by deriving three general convergence results towards consensus depending on the nature of the interaction topology, and presented in ascending order of generality and difficulty. We thus start in Section 4.2.1 by the simplest scenario in which the adjacency operators  $\mathcal{A}(t) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$ , whose kernels are  $a(t) \in L^\infty(I \times I, [0, 1])$ , define symmetric interaction topologies for almost every time  $t \geq 0$ . By adapting a methodology developed in [14] by the first author, we are then able to prove in Theorem 4.8 that solutions of (1.7) exponentially converge to consensus in the  $L^2$ -norm topology, whenever the interaction topology satisfies the persistence condition

$$\lambda_2\left(\frac{1}{\tau} \int_t^{t+\tau} \mathbb{L}^a(s) ds\right) \geq \mu \quad (1.9)$$

for all times  $t \geq 0$ , in which  $\mathbb{L}^a(t)$  stands for the linear graph-Laplacian

$$\mathbb{L}^a(t) : x \mapsto \int_I a(t, i, j)(x(i) - x(j)) dj,$$

generated by the adjacency operator  $\mathcal{A}(t)$  at some time  $t \geq 0$ . It is worth recalling that, for finite-dimensional multi-agent systems, persistence conditions of the form (1.9) bearing on the smallest eigenvalue of the averaged matrix are in fact *necessary and sufficient* for the exponential convergence towards consensus, as shown in [51]. We subsequently move on in Section 4.2.2 to the more general situation in which the interaction graph is directed, but balanced, in the sense that the in- and out-degrees coincide at each node. Contrary to the symmetric case, we need to restrict our attention to linear dynamics in this context. In Theorem 4.11, we show that solutions of (1.7) for which  $\phi \equiv 1$  and  $\mathcal{A}(t)$  defines a balanced topology for almost all  $t \geq 0$  exponentially converge to consensus in the  $L^2$ -norm topology, whenever

$$\frac{1}{\tau} \int_t^{t+\tau} \lambda_2(\mathbb{L}(s)) ds \geq \mu \quad (1.10)$$

for all times  $t \geq 0$ . At this point, a crucial observation has to be made regarding the persistence conditions (1.9) and (1.10). Indeed, the latter – which heuristically imposes that the algebraic connectivity of the system be positive sufficiently often – is much more restrictive than the former, which merely requires that the average graph on every time window of length  $\tau > 0$  be strongly connected. This requirement can be met even in the extreme case where  $\lambda_2(\mathbb{L}^a(t)) = 0$  for almost every times  $t \geq 0$ , which corresponds to the scenario in which the interaction graph is disconnected at all times. This intuition is made rigorous in Proposition 4.12, where we show that  $\lambda_2(\cdot)$  is in fact a concave mapping over the space of linear operators. Finally in Section 4.2.3, we treat the more general situation in which the adjacency operators  $\mathcal{A}(t)$  define topologies which are disjoint unions of strongly connected components – see Definition 4.5 below – for almost every time  $t \geq 0$ . These are much more delicate to handle than balanced topologies, as the system admits dissipative Lyapunov functions which are different for each possible interaction topology. For this reason, even though the corresponding exponential convergence result is supported by a suitable generalisation of the persistence condition (1.10), we need to impose additional assumptions on the structure of the interactions to be able to compare

the different energy functionals. The corresponding result, which builds on the methodology introduced for balanced graphons combined with stability techniques for switching systems, is presented in Theorem 4.13.

We conclude the paper in Section 5 by establishing several auxiliary results which shed light on the interplay between consensus formation and graphon connectivity. In Section 5.1, we generalise to symmetric graphon models one of the instrumental results about algebraic connectivity, namely that the underlying interaction topology is strongly connected if and only if  $\lambda_2(\mathbb{L}) > 0$ . We then prove in Section 5.2 a somewhat surprising result, stating that  $L^2$ - and  $L^\infty$ -consensus formation are in fact equivalent whenever the in-degree function associated with the graphon model is persistent in a suitable sense. We end the article by displaying numerical illustrations of our consensus results in Section 5.3, along with a discussion on some of the limit cases of the theory that we develop here.

The structure of the article is the following. In Section 2, we present standard notions of measure, graphs and spectral theory, and provide a short primer on discrete multi-agent systems and graphon models. In particular, we prove in this occasion that the convex hull of a graphon dynamics is non-increasing, which is a new and interesting result in its own right. In Section 3, we define scrambling coefficients for graphons, and study diameter contraction estimates and  $L^\infty$ -consensus formation. In Section 4, we adopt another point of view and study the formation of  $L^2$ -consensus for symmetric, balanced, and strongly connected interaction topologies. These latter are all based on adequate generalisations of the notion of algebraic connectivity for graph-Laplacian operators. We finally conclude the paper by proving auxiliary results and presenting numerical illustrations in Section 5.

## 2 Preliminaries

In this section, we recollect a series of notions pertaining to functional analysis and multi-agent dynamics at large, all of which will be extensively used throughout the article.

### 2.1 Measure theory and integration

We start here by introducing some notations and results of measure theory, for which we refer the reader to the very complete monographs [1, 33].

In what follows, given an integer  $n \geq 1$ , we denote by  $\mathcal{L}^n$  the standard  $n$ -dimensional Lebesgue measure defined over  $\mathbb{R}^n$ . Letting  $p \in [1, +\infty]$ , taking a Borel set  $\Omega \subset \mathbb{R}^n$  and denoting by  $(X, \|\cdot\|_X)$  a separable Banach space, we shall use the notation  $L^p(\Omega, X)$  for the space of  $X$ -valued integrable maps with respect to  $\mathcal{L}^n$ , defined in the sense of Bochner (see e.g. [31, Chapter II]). We will also write  $C^0(\Omega, \mathcal{S})$  and  $\text{Lip}_{\text{loc}}(\Omega, \mathcal{S})$  to refer to continuous and locally Lipschitz mappings with values in a complete separable metric space  $(\mathcal{S}, d_{\mathcal{S}})$  respectively. Throughout this article, we will use the notations “inf”, “sup” and “rg” to refer to the essential infimum, supremum and range of a measurable function, all of which are understood with respect to the standard Lebesgue measure.

We state below simplified versions of the *Scorza-Dragoni* and *Danskin* theorems (see [8] and [30] respectively), which will be used in the sequel. While these latter are both known results in modern analysis, they are somewhat nonstandard, and we include their statements for the sake of completeness.

**Theorem 2.1** (Scorza-Dragoni). *Let  $\Omega \subset \mathbb{R}^n$  be a Borel set and  $f : \mathbb{R}_+ \times \Omega \rightarrow \mathcal{S}$  be such that  $x \in \Omega \mapsto f(t, x) \in \mathcal{S}$  is  $\mathcal{L}^n$ -measurable for each  $t \geq 0$ , and  $t \in \mathbb{R}_+ \mapsto f(t, x) \in \mathcal{S}$  is continuous for  $\mathcal{L}^n$ -almost every  $x \in \Omega$ . Then for every  $\varepsilon > 0$ , there exists a compact set  $\Omega_\varepsilon \subset \Omega$  satisfying  $\mathcal{L}^n(\Omega \setminus \Omega_\varepsilon) < \varepsilon$  and such that the restricted map  $f : \mathbb{R}_+ \times \Omega_\varepsilon \rightarrow \mathcal{S}$  is continuous.*

**Theorem 2.2** (Danskin). *Let  $\Omega \subset \mathbb{R}^n$  be a compact set and  $f : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}$  be a continuous function such that  $t \in \mathbb{R}_+ \mapsto f(t, x) \in \mathbb{R}$  is differentiable for all  $x \in \Omega$ . Then, the application  $g : t \in \mathbb{R}_+ \mapsto \max_{x \in \Omega} f(t, x) \in \mathbb{R}$  is differentiable  $\mathcal{L}^1$ -almost everywhere, with*

$$\frac{d}{dt}g(t) = \max_{x \in \bar{\Omega}(t)} \partial_t f(t, x)$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ , where we introduced the notation  $\bar{\Omega}(t) := \operatorname{argmax}_{x \in \Omega} f(t, x)$ .

In the remainder of this article, we will often use both theorems in conjunction with the following convergence result for restricted essential supremums. Therein, we denote by  $\mathcal{P}(\Omega)$  the power set generated by some arbitrary  $\Omega \subset \mathbb{R}^n$ .

**Lemma 2.3** (Quantitative interior estimates for essential supremums). *Let  $\Omega \subset \mathbb{R}^n$  be a compact set and  $f \in L^\infty(\Omega, \mathbb{R}^d)$ . Then for every  $\delta > 0$ , there exists  $\varepsilon > 0$  such that*

$$\|f\|_{L^\infty(\Omega, \mathbb{R}^d)} - \delta \leq \|f\|_{L^\infty(\Omega_\varepsilon, \mathbb{R}^d)} \leq \|f\|_{L^\infty(\Omega, \mathbb{R}^d)}, \quad (2.1)$$

whenever  $\Omega_\varepsilon \subset \Omega$  is a measurable set satisfying  $\mathcal{L}^n(\Omega \setminus \Omega_\varepsilon) < \varepsilon$ . In particular, it holds that

$$\|f\|_{L^\infty(\Omega_\varepsilon, \mathbb{R}^d)} \xrightarrow{\varepsilon \rightarrow 0^+} \|f\|_{L^\infty(\Omega, \mathbb{R}^d)},$$

for each family of sets  $(\Omega_\varepsilon)_{\varepsilon > 0} \subset \mathcal{P}(\Omega)$  satisfying these properties.

*Proof.* The second inequality in (2.1) is trivially satisfied for every set  $\Omega_\varepsilon \subset \Omega$ . Concerning the first inequality, observe that by definition of the  $L^\infty$ -norm, there exists for every  $\delta > 0$  a set  $\Omega_\delta \subset \Omega$  of positive measure such that

$$|f(x)| \geq \|f\|_{L^\infty(\Omega, \mathbb{R}^d)} - \delta$$

for  $\mathcal{L}^n$ -almost every  $x \in \Omega_\delta$ . Hence, by choosing  $\varepsilon > 0$  satisfying  $\varepsilon < \mathcal{L}^n(\Omega_\delta)$  and considering any closed set  $\Omega_\varepsilon \subset \Omega$  with  $\mathcal{L}^n(\Omega \setminus \Omega_\varepsilon) < \varepsilon$ , it necessarily holds that  $\mathcal{L}^n(\Omega_\delta \cap \Omega_\varepsilon) > 0$ . Indeed assuming the converse by contradiction, one would have

$$\mathcal{L}^n(\Omega_\delta \cup \Omega_\varepsilon) = \mathcal{L}^n(\Omega_\delta) + \mathcal{L}^n(\Omega_\varepsilon) > \mathcal{L}^n(\Omega),$$

which is absurd since both sets are contained in  $\Omega$ . Thus by construction, one has that

$$|f(x)| \geq \|f\|_{L^\infty(\Omega, \mathbb{R}^d)} - \delta,$$

for  $\mathcal{L}^n$ -almost every  $x \in \Omega_\delta \cap \Omega_\varepsilon$ , which concludes the proof of our claim.  $\square$

## 2.2 Graph and spectral theory

In this second preliminary section, we recollect more specifically elementary notions pertaining to graph theory, together with elements of spectral analysis for bounded linear operators defined over Hilbert spaces. We point to the reference monographs [78] and [69] respectively for a detailed and accessible treatment of these topics.

**Introductory notions of graph theory.** Given an integer  $N \geq 1$ , we say that  $\mathbf{A}_N := (a_{ij})_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times N}$  is the *adjacency matrix* of a directed graph – or digraph – with vertex labels  $\{1, \dots, N\}$  provided that  $a_{ij} \in [0, 1]$  and  $a_{ii} = 1$  for each  $i, j \in \{1, \dots, N\}$ . In the context of cooperative dynamics, the set of labels will represent the agents of the system, while each coefficient  $a_{ij} \in [0, 1]$  stands for the magnitude of communication flowing from agent  $j$  to agent  $i$ . In particular,  $a_{ij} > 0$  if and only if agent  $j$  influences the dynamics of agent  $i$ . We recall below standard notions of connectivity for general digraphs, which are illustrated in Figure 2.1

**Definition 2.4** (Strong and simple digraph connectivity). *Let  $\mathbf{A}_N \in [0, 1]^{N \times N}$  be an adjacency matrix. Then, the underlying digraph is strongly connected if for every pair of indices  $i, j \in \{1, \dots, N\}$ , there exist an integer  $m \geq 1$  and a finite sequence  $(l_k)_{1 \leq k \leq m} \subset \{1, \dots, N\}$  such that  $i = l_1$ ,  $j = l_m$  and  $a_{l_k l_{k+1}} > 0$  for each  $k \in \{1, \dots, m-1\}$ .*

*In addition, a digraph is said to be simply connected if its symmetric part – which is defined as the undirected graph supported by the adjacency matrix  $\frac{1}{2}(\mathbf{A}_N + \mathbf{A}_N^\top)$  –, is strongly connected. In particular, an undirected graph is connected if and only if it is strongly connected.*



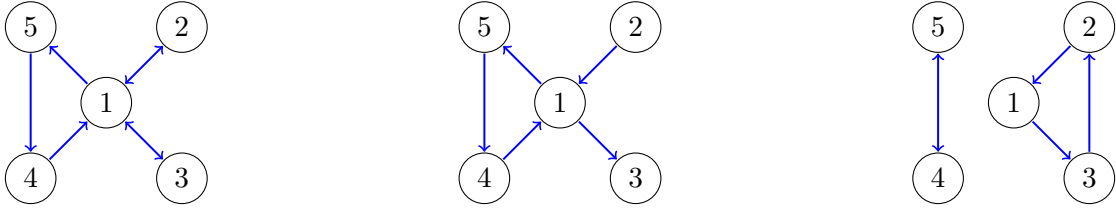


Figure 2.1: Illustration of the notions of strong and simple connectivity on graphs with  $N = 5$  vertices. The digraph on the left is strongly connected, while that in the center is simply connected. The digraph on the right is the disjoint union of two strongly connected components formed respectively by  $\{1, 2, 3\}$  and  $\{4, 5\}$ .

**Elements of spectral theory for bounded operators.** Let  $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$  be an arbitrary finite- or infinite-dimensional Hilbert space, and  $\mathcal{L}(\mathcal{H})$  denote the space of all bounded linear operators from  $\mathcal{H}$  into itself. We will say that an element  $T \in \mathcal{L}(\mathcal{H})$  is *symmetric* if  $T = T^*$ , where  $T^* \in \mathcal{L}(\mathcal{H})$  is the adjoint of  $T$  with respect to  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ . We also use the notation  $T^{\text{sym}} := \frac{1}{2}(T + T^*)$  for its symmetric part. Notice that a symmetric operator which is bounded – namely whose domain is the whole space – is automatically self-adjoint in the usual sense. Writing  $\text{Id}$  for the identity operator over  $\mathcal{H}$ , we recall below several concepts pertaining to the *spectrum* of an element of  $\mathcal{L}(\mathcal{H})$ .

**Definition 2.5** (The spectrum of bounded linear operators and its decomposition). *Given an element  $T \in \mathcal{L}(\mathcal{H})$ , its spectrum  $\sigma(T)$  is defined as the set of all the real numbers  $\lambda \in \mathbb{R}$  such that*

$$(T - \lambda \text{Id}) \in \mathcal{L}(\mathcal{H}) \quad \text{is not bijective.}$$

*In addition, the spectrum of a bounded linear operator  $T$  admits the following decomposition*

$$\sigma(T) = \sigma_{\text{disc}}(T) \cup \sigma_{\text{ess}}(T),$$

*where the union is disjoint. Therein,  $\sigma_{\text{disc}}(T)$  denotes the discrete spectrum of  $T$ , defined by*

$$\sigma_{\text{disc}}(T) := \left\{ \lambda \in \sigma(T) \text{ s.t. there exists an open set } \mathcal{N}_{\lambda} \subset \mathbb{R} \text{ satisfying } \mathcal{N}_{\lambda} \cap \sigma(T) = \{\lambda\} \right. \\ \left. \text{and the space of all } f \in \mathcal{H} \text{ such that } Tf = \lambda f \text{ is finite-dimensional} \right\},$$

*while  $\sigma_{\text{ess}}(T) := \sigma(T) \setminus \sigma_{\text{disc}}(T)$  is the essential spectrum of  $T$ .*

We will use the notation  $I := [0, 1]$  for the unit interval representing the continuum of indices, and denote by  $\mathcal{L}_{\square I}^1$  the restriction of the standard 1-dimensional Lebesgue measure thereon. We shall say that  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is an *adjacency operator* if there exists a map  $a \in L^\infty(I \times I, [0, 1])$  such that

$$(\mathcal{A}x)(i) := \int_I a(i, j)x(j)dj \quad \text{for } \mathcal{L}_{\square I}^1\text{-almost every } i \in I, \quad (2.2)$$

for each  $x \in L^2(I, \mathbb{R}^d)$ . Analogously given a map  $d \in L^\infty(I, [0, 1])$ , we define the corresponding *multiplication operator*  $\mathcal{M}_d \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  by

$$(\mathcal{M}_d x)(i) := d(i)x(i) \quad \text{for } \mathcal{L}_{\square I}^1\text{-almost every } i \in I, \quad (2.3)$$

for each  $x \in L^2(I, \mathbb{R}^d)$ . We end this preliminary section by recalling the following standard structure result on the spectra of adjacency and multiplication operators, for which we point to [69, Theorem VI.15] and [40, Proposition 3.2] respectively.

**Proposition 2.6** (Spectra of adjacency and multiplication operators). *Let  $\mathcal{A}, \mathcal{M}_d \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  be an adjacency and a multiplication operator defined as in (2.2) and (2.3) respectively. Then, one has*

$$\sigma(\mathcal{A}) = \sigma_{\text{disc}}(\mathcal{A}) = \{\lambda_n\}_{n=1}^{+\infty} \quad \text{and} \quad \sigma(\mathcal{M}_d) = \sigma_{\text{ess}}(\mathcal{M}_d) = \text{rg}(d),$$

*where  $(\lambda_n) \subset [0, 1]$  is an at most countable family of real numbers which can only accumulate at zero.*

## 2.3 Multi-agent systems and graphon models

In this section, we recall some of the properties of the multi-agent systems under consideration throughout this article, both at the discrete and macroscopic level. We point the reader to [24, 62] for a detailed account on mean-field limits for particle systems, and to [49, 55] for a thorough analysis of graph limits.

**Finite-dimensional cooperative systems.** We start by considering a system of  $N \geq 1$  agents, represented by a collection  $\mathbf{x}(\cdot) := (x_1(\cdot), \dots, x_N(\cdot)) \in C^0(\mathbb{R}_+, (\mathbb{R}^d)^N)$  of curves in  $\mathbb{R}^d$ . Given a set  $(x_1^0, \dots, x_N^0) \in (\mathbb{R}^d)^N$  of initial data, their evolution in time is prescribed by the family of ODEs

$$\begin{cases} \dot{x}_i(t) = \frac{1}{N} \sum_{j=1}^N a_{ij}(t) \phi(|x_i(t) - x_j(t)|) (x_j(t) - x_i(t)), \\ x_i(0) = x_i^0, \end{cases} \quad (2.4)$$

written for every  $i \in \{1, \dots, N\}$ . Here, the positive nonlinear interaction kernel  $\phi(\cdot)$  accounts for the magnitude of the pairwise interactions between agents – depending on their relative distance –, while the maps  $(a_{ij}(\cdot))_{1 \leq i, j \leq N}$  encode the time-varying interaction topology of the system.

In the sequel, we make the following assumptions on the data of (2.4).

**Hypotheses (MA).** *Assume that the following holds.*

- (i) *The weight functions  $a_{ij} : \mathbb{R}_+ \rightarrow [0, 1]$  are  $\mathcal{L}^1$ -measurable for all  $i, j \in \{1, \dots, N\}$ . Moreover, one has that  $a_{ii}(t) = 1$  for  $\mathcal{L}^1$ -almost every  $t \geq 0$  and each  $i \in \{1, \dots, N\}$ .*
- (ii) *The nonlinear kernel  $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+^*$  is locally Lipschitz continuous, positive and bounded.*

Let it be noted that the weight functions  $a_{ij}(\cdot)$  could more generally be allowed to take any positive value as long as they are bounded, but that the corresponding range can always be normalised to  $[0, 1]$  e.g. by suitably rescaling the nonlinear kernel  $\phi(\cdot)$ . Under this series of assumptions – which are very standard in the literature devoted to consensus problems, see e.g. [20, 61] –, we have the following well-posedness result for (2.4).

**Proposition 2.7** (Well-posedness of the discrete dynamics). *Let  $\mathbf{x}^0 := (x_1^0, \dots, x_N^0)$  and suppose that hypotheses (MA) hold. Then, there exists a unique curve  $\mathbf{x} \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, (\mathbb{R}^d)^N)$  solution of (2.4).*

*Proof.* It can be easily checked that under hypotheses (MA), the mapping

$$(t, \mathbf{x}) \in \mathbb{R}_+ \times (\mathbb{R}^d)^N \mapsto \left( \frac{1}{N} \sum_{j=1}^N a_{ij}(t) \phi(|x_i - x_j|) (x_j - x_i) \right)_{1 \leq i \leq N} \in (\mathbb{R}^d)^N,$$

is  $\mathcal{L}^1$ -measurable with respect to  $t \geq 0$  as well as locally Lipschitz continuous and sublinear with respect to  $\mathbf{x} \in (\mathbb{R}^d)^N$ . Whence, by classical well-posedness results for ordinary differential equations (see e.g. [35, Chapter 1]), there exists a unique solution  $\mathbf{x}(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, (\mathbb{R}^d)^N)$  to (2.4).  $\square$

**Macroscopic approximations of multi-agent systems.** In the classical mean-field setting, macroscopic versions of the discrete dynamics (2.4) are expressed by curves of empirical measures. However, as explained in the Introduction, this approach is only operational when the agents are *indistinguishable*, or equivalently provided that the dynamics (2.4) satisfies some invariance properties with respect to the agent labels  $i \in \{1, \dots, N\}$ . To circumvent this intrinsic limitation, one can instead consider *graph limit* approximations, which are built as follows. Consider the piecewise constant mappings

$$x^N(t, i) := \sum_{k=1}^N \mathbb{1}_{\left[\frac{k-1}{N}, \frac{k}{N}\right)}(i) x_k(t) \quad (2.5)$$

and

$$a^N(t, i, j) := \sum_{k=1}^N \sum_{l=1}^N \mathbb{1}_{\left[\frac{k-1}{N}, \frac{k}{N}\right)}(i) \mathbb{1}_{\left[\frac{l-1}{N}, \frac{l}{N}\right)}(j) a_{kl}(t), \quad (2.6)$$

defined for  $\mathcal{L}^1$ -almost every  $t \geq 0$  and every  $i, j \in I$ . Then, one can show that the curve  $x^N(\cdot) \in C^0(\mathbb{R}_+, L^2(I, \mathbb{R}^d))$  satisfies the following.

**Proposition 2.8** (Dynamics of the discrete graphon model). *Let  $N \geq 1$  and  $\mathbf{x}(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, (\mathbb{R}^d)^N)$  be a solution of (2.4). Then, the curve  $x^N(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, L^2(I, \mathbb{R}^d))$  defined in (2.5) is a solution of the Cauchy problem*

$$\begin{cases} \partial_t x^N(t, i) = \int_I a^N(t, i, j) \phi(|x^N(t, i) - x^N(t, j)|) (x^N(t, j) - x^N(t, i)) \text{d}j, \\ x^N(0, i) = x^0(i), \end{cases} \quad (2.7)$$

formulated for  $\mathcal{L}^1_I$ -almost every  $i \in I$ , where  $a^N \in L^\infty(\mathbb{R}_+ \times I \times I, [0, 1])$  is defined as in (2.6).

*Proof.* Observe first that, for  $\mathcal{L}^1_I$ -almost every  $i \in I$ , the map  $t \in \mathbb{R}_+ \mapsto x^N(t, i) \in \mathbb{R}^d$  is differentiable  $\mathcal{L}^1$ -almost everywhere. Then, by combining (2.4) together with (2.5), one has that

$$\begin{aligned} \partial_t x^N(t, i) &= \sum_{k=1}^N \mathbf{1}_{\left[\frac{k-1}{N}, \frac{k}{N}\right)}(i) \dot{x}_k(t) \\ &= \sum_{k=1}^N \mathbf{1}_{\left[\frac{k-1}{N}, \frac{k}{N}\right)}(i) \left( \frac{1}{N} \sum_{l=1}^N a_{kl}(t) \phi(|x_k(t) - x_l(t)|) (x_l(t) - x_k(t)) \right) \\ &= \sum_{l=1}^N \int_{(l-1)/N}^{l/N} \left( \sum_{k=1}^N \mathbf{1}_{\left[\frac{k-1}{N}, \frac{k}{N}\right)}(i) a_{kl}(t) \phi(|x_k(t) - x_l(t)|) (x_l(t) - x_k(t)) \right) \text{d}j \\ &= \int_I a^N(t, i, j) \phi(|x^N(t, i) - x^N(t, j)|) (x^N(t, j) - x^N(t, i)) \text{d}j, \end{aligned}$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$  and  $\mathcal{L}^1_I$ -almost every  $i \in I$ , which concludes the proof of our claim.  $\square$

**Remark 2.9** (Subordination of finite multi-agent systems to graphon models). *It is worth noting that any finite-dimensional multi-agent system of the form (2.4) can be recast as a graphon dynamics via (2.5) and (2.6). This implies in particular that all the results that are discussed hereinbelow for general graphon models – including the well-posedness results and sufficient conditions for various types of consensus formation – have suitable counterparts in the discrete setting.*

**Well-posedness and elementary properties of graphon models.** While the dynamics in (2.7) describes the evolution of a piecewise constant macroscopic approximation of (2.4), it can more generally be seen as an infinite-dimensional integro-differential equation. Therefore in the sequel, given an initial datum  $x^0 \in L^2(I, \mathbb{R}^d)$ , we shall study the well-posedness of *graphon dynamics* of the form

$$\begin{cases} \partial_t x(t, i) = \int_I a(t, i, j) \phi(|x(t, i) - x(t, j)|) (x(t, j) - x(t, i)) \text{d}j, \\ x(0) = x^0. \end{cases} \quad (2.8)$$

In this context, the time-varying interaction topology is encoded by a measurable mapping  $a \in L^\infty(\mathbb{R}_+ \times I \times I, [0, 1])$ , while the distance-based communications between agents remain described by a nonlinear interaction kernel  $\phi(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, \mathbb{R}_+^*)$ .

Throughout this manuscript, we will make the following assumptions on the data of the graphon dynamics (2.8).

**Hypotheses (GD).** *Assume that the following holds.*

- (i) *The weight function  $a : \mathbb{R}_+ \times I \times I \rightarrow [0, 1]$  is  $\mathcal{L}^1 \times \mathcal{L}^1_I \times \mathcal{L}^1_I$ -measurable.*
- (ii) *The nonlinear kernel  $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+^*$  is locally Lipschitz continuous, positive and bounded.*

We refer the reader to [4] for a detailed and pedagogical introduction to graphon models in the context of multi-agent systems.

**Proposition 2.10** (Well-posedness and estimates for graphon dynamics). *Let  $x^0 \in L^2(I, \mathbb{R}^d)$  and assume that hypotheses **(GD)** hold. Then, there exists a unique solution  $x(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, L^2(I, \mathbb{R}^d))$  to the Cauchy problem (2.8). If in addition it holds that  $x^0 \in L^\infty(I, \mathbb{R}^d)$ , one then has*

$$\|x(t)\|_{L^\infty(I, \mathbb{R}^d)} \leq \|x^0\|_{L^\infty(I, \mathbb{R}^d)}, \quad (2.9)$$

for all times  $t \geq 0$ .

*Proof.* The existence and uniqueness of a solution  $x(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, L^2(I, \mathbb{R}^d))$  to (2.8) can be recovered by repeating the method of [4, Section 3.1], up to some minor modifications. Assume now that  $x^0 \in L^\infty(I, \mathbb{R}^d)$ , and observe that by integrating (2.8) with respect to  $t \geq 0$ , one has

$$\begin{aligned} |x(t, i)| &\leq |x^0(i)| + \int_0^t \int_I a(s, i, j) \phi(|x(s, i) - x(s, j)|) |x(s, j) - x(s, i)| \, dj \, ds \\ &\leq \|x^0\|_{L^\infty(I, \mathbb{R}^d)} + c_\phi \int_0^t \left( |x(s, i)| + \int_I |x(s, j)| \, dj \right) \, ds, \end{aligned} \quad (2.10)$$

for all times  $t \geq 0$  and  $\mathcal{L}_{\perp I}^1$ -almost every  $i \in I$ , where  $c_\phi := \sup_{r \in \mathbb{R}_+} \phi(r) < +\infty$ . By integrating (2.10) with respect to  $i \in I$  and applying Fubini's theorem along with Grönwall's lemma, we obtain

$$\int_I |x(t, i)| \, di \leq \|x^0\|_{L^\infty(I, \mathbb{R}^d)} \exp(2c_\phi t), \quad (2.11)$$

for all times  $t \geq 0$ . Then, plugging (2.11) into (2.10) while resorting again to Grönwall's lemma yields

$$|x(t, i)| \leq \|x^0\|_{L^\infty(I, \mathbb{R}^d)} \left( 1 + \exp(2c_\phi t) \right) \frac{\exp(c_\phi t)}{2}, \quad (2.12)$$

for  $\mathcal{L}_{\perp I}^1$ -almost every  $i \in I$ , which implies that  $x(t) \in L^\infty(I, \mathbb{R}^d)$  for all times  $t \geq 0$ .

Our goal now is to prove the sharper stability estimate displayed (2.9). To this end, fix a real number  $\varepsilon > 0$ , and observe that by Scorza-Dragoni's theorem (see Theorem 2.1 above), there exists a compact subset of indices  $I_\varepsilon \subset I$  satisfying  $\mathcal{L}_{\perp I}^1(I \setminus I_\varepsilon) < \varepsilon$  such that the mapping  $x : \mathbb{R}_+ \times I_\varepsilon \rightarrow \mathbb{R}^d$  is continuous. We then define for all times  $t \geq 0$  the family of restricted supremum norms

$$L_\varepsilon(t) := \max_{i \in I_\varepsilon} |x(t, i)|.$$

The mapping  $L_\varepsilon(\cdot)$  is locally Lipschitz as the pointwise maximum of a family of equi-locally Lipschitz functions, and is therefore differentiable  $\mathcal{L}^1$ -almost everywhere by Rademacher's theorem (see e.g. [33, Theorem 3.2]). Moreover by Danskin's theorem (see Theorem 2.2 above), one has that

$$\frac{1}{2} \frac{d}{dt} L_\varepsilon(t)^2 = \max_{i \in J_\varepsilon(t)} \langle \partial_t x(t, i), x(t, i) \rangle, \quad (2.13)$$

where  $J_\varepsilon(t) := \text{argmax}_{i \in I_\varepsilon} |x(t, i)|$  for  $\mathcal{L}^1$ -almost every  $t \geq 0$ . Taking an arbitrary  $i \in J_\varepsilon(t)$ , one observes that

$$\begin{aligned} \langle \partial_t x(t, i), x(t, i) \rangle &= \int_I a(t, i, j) \phi(|x(t, i) - x(t, j)|) \langle x(t, i), x(t, j) - x(t, i) \rangle \, dj \\ &\leq \int_{I_\varepsilon} a(t, i, j) \phi(|x(t, i) - x(t, j)|) L_\varepsilon(t) (|x(t, j)| - L_\varepsilon(t)) \, dj \\ &\quad + \int_{I \setminus I_\varepsilon} a(t, i, j) \phi(|x(t, i) - x(t, j)|) L_\varepsilon(t) (|x(t, j)| - L_\varepsilon(t)) \, dj \\ &\leq c_\phi \varepsilon \|x(t)\|_{L^\infty(I, \mathbb{R}^d)}^2, \end{aligned} \quad (2.14)$$

where we used the facts that  $|x(t, j)| \leq L_\varepsilon(t)$  for each  $j \in I_\varepsilon$  and  $\mathcal{L}_{\perp I}^1(I \setminus I_\varepsilon) < \varepsilon$  by construction. By merging (2.13) together with (2.14) and applying Grönwall's lemma, we further obtain the estimate

$$\frac{1}{2} L_\varepsilon(t)^2 \leq \frac{1}{2} L_\varepsilon(0)^2 + c_\phi \varepsilon \int_0^t \|x(s)\|_{L^\infty(I, \mathbb{R}^d)}^2 \, ds, \quad (2.15)$$

which holds for all times  $t \geq 0$ . Observing that by Lemma 2.3, we have the pointwise convergence

$$L_\varepsilon(t) \xrightarrow{\varepsilon \rightarrow 0^+} \|x(t)\|_{L^\infty(\Omega, \mathbb{R}^n)},$$

we can pass to the limit as  $\varepsilon \rightarrow 0^+$  in (2.15) and conclude that

$$\|x(t)\|_{L^\infty(I, \mathbb{R}^d)} \leq \|x^0\|_{L^\infty(I, \mathbb{R}^d)},$$

for all times  $t \geq 0$ , which ends the proof of our claim.  $\square$

Another feature of finite-dimensional multi-agent systems that still holds for graphon models is that the *closed convex hull* of the essential range of the initial datum  $x^0 \in L^\infty(I, \mathbb{R}^d)$ , defined by

$$\mathcal{C}(x^0) = \overline{\text{co}}(\text{rg}(x^0)) := \overline{\bigcap \left\{ K \subset \mathbb{R}^d \text{ s.t. } K \subset \mathbb{R}^d \text{ is convex and } \text{rg}(x^0) \subset K \right\}},$$

is invariant under the dynamics (2.8), as shown in the following proposition.

**Proposition 2.11** (Invariance of the convex hull). *Let  $x^0 \in L^\infty(I, \mathbb{R}^d)$  and  $x(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, L^2(I, \mathbb{R}^d))$  be the corresponding solution of the graphon dynamics (2.8). Then, the closed convex hulls satisfy the inclusion  $\mathcal{C}(x(t)) \subset \mathcal{C}(x^0)$  for all times  $t \geq 0$ .*

*Proof.* Similarly to what was done in the proof of Proposition 2.10, we start by fixing  $\varepsilon > 0$  and let  $I_\varepsilon \subset I$  be a compact subset given by Scorza-Dragnoni's theorem (see Theorem 2.1 above). We also consider the restricted closed convex hulls, defined by

$$\mathcal{C}_\varepsilon(x(t)) := \bigcap \left\{ K \subset \mathbb{R}^d \text{ s.t. } K \subset \mathbb{R}^d \text{ is convex and } x(t, I_\varepsilon) \subset K \right\}$$

for all times  $t \geq 0$ , which are compact sets because  $x(t, I_\varepsilon)$  is compact. Furthermore, recall that as a consequence of the separation theorem (see e.g. [3, Theorem 2.4.2]), the closed convex hull of the compact sets  $x(t, I_\varepsilon) \subset \mathbb{R}^d$  and  $\text{rg}(x(t))$  can be characterised respectively by the identities

$$\mathcal{C}_\varepsilon(x(t)) = \left\{ z \in \mathbb{R}^d \text{ s.t. } \langle p, z \rangle \leq \max_{i \in I_\varepsilon} \langle p, x(t, i) \rangle \text{ for each } p \in \mathbb{R}^d \right\}, \quad (2.16)$$

and

$$\mathcal{C}(x(t)) = \left\{ z \in \mathbb{R}^d \text{ s.t. } \langle p, z \rangle \leq \sup_{i \in I} \langle p, x(t, i) \rangle \text{ for each } p \in \mathbb{R}^d \right\}, \quad (2.17)$$

where we used the fact that the essential supremum of a real-valued measurable map coincides with the maximum value of its essential range.

Given an element  $p \in \mathbb{R}^d$ , set  $J_{\varepsilon, p}(t) := \text{argmax}_{i \in I_\varepsilon} \langle p, x(t, i) \rangle$  and observe that by Danskin's theorem (see Theorem 2.2 above), it holds that

$$\frac{d}{dt} \max_{i \in I_\varepsilon} \langle p, x(t, i) \rangle = \max_{i \in J_{\varepsilon, p}(t)} \langle p, \partial_t x(t, i) \rangle,$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ . For an arbitrary  $i \in J_{\varepsilon, p}(t)$ , one further has

$$\begin{aligned} \langle p, \partial_t x(t, i) \rangle &= \int_I a(t, i, j) \phi(|x(t, i) - x(t, j)|) \langle p, x(t, j) - x(t, i) \rangle dj \\ &= \int_{I_\varepsilon} a(t, i, j) \phi(|x(t, i) - x(t, j)|) \langle p, x(t, j) - x(t, i) \rangle dj + \varepsilon C_p(t), \end{aligned} \quad (2.18)$$

where  $C_p(\cdot) \in L^\infty_{\text{loc}}(\mathbb{R}_+, \mathbb{R})$ . Notice now that the under hypotheses **(GD)**, the identity

$$\int_{I_\varepsilon} a(t, i, j) \phi(|x(t, i) - x(t, j)|) dj = 0, \quad (2.19)$$

can only hold true provided that  $a(t, i, j) = 0$  for  $\mathcal{L}^1_I$ -almost every  $j \in I_\varepsilon$ . In that case, (2.18) can be simply rewritten as

$$\langle p, \partial_t x(t, i) \rangle = \varepsilon C_p(t), \quad (2.20)$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ . On the contrary, upon assuming that (2.19) does not hold, one then has

$$\langle p, \partial_t x(t, i) \rangle = \left( \int_{I_\varepsilon} a(t, i, j) \phi(|x(t, i) - x(t, j)|) dj \right) \langle p, \bar{X}_\varepsilon(t, i) - x(t, i) \rangle + \varepsilon C_p(t), \quad (2.21)$$

where the weighted barycenter point, defined by

$$\bar{X}_\varepsilon(t, i) = \int_{I_\varepsilon} \left( \frac{a(t, i, j) \phi(|x(t, i) - x(t, j)|) x(t, j)}{\int_{I_\varepsilon} a(t, i, k) \phi(|x(t, i) - x(t, k)|) dk} \right) dj, \quad (2.22)$$

is an element of  $\mathcal{C}_\varepsilon(x(t))$  by the standard convexification property of the integral (see e.g. [2, Chapter 5, Theorem 3]). Whence, using the characterisation (2.16) of the convex hull together with the definition of the set of indices  $J_{\varepsilon, p}(t)$  in the identity (2.21) finally yields

$$\langle p, \partial_t x(t, i) \rangle \leq \varepsilon C_p(t), \quad (2.23)$$

for all times  $t \geq 0$  and any  $i \in J_{\varepsilon, p}(t)$ . Thus by combining (2.20) and (2.23) and performing an integration with respect to the time variable, we recover the pointwise inequality

$$\max_{i \in I_\varepsilon} \langle p, x(t, i) \rangle \leq \max_{i \in I_\varepsilon} \langle p, x^0(i) \rangle + \varepsilon \int_0^t C_p(s) ds,$$

for all times  $t \geq 0$ . By letting  $\varepsilon \rightarrow 0^+$  and applying the convergence result of Lemma 2.3, it thus holds

$$\sup_{i \in I} \langle p, x(t, i) \rangle \leq \sup_{i \in I} \langle p, x^0(i) \rangle,$$

and in particular, by choosing an element  $y \in \mathcal{C}(x(t))$ , one has that

$$\langle p, y \rangle \leq \sup_{i \in I} \langle p, x(t, i) \rangle \leq \sup_{i \in I} \langle p, x^0(i) \rangle,$$

for every  $p \in \mathbb{R}^d$  as a consequence of (2.17). We have thus shown that  $\mathcal{C}(x(t)) \subset \mathcal{C}(x^0)$  for all times  $t \geq 0$ , which concludes the proof of our claim.  $\square$

**Remark 2.12** (Performing differential estimates combining Scorza-Dragoni and Danskin theorems). *In the setting of discrete multi-agent models, the essential suprema over  $I = [0, 1]$  reduce to simple maxima of a finite collection of equi-locally Lipschitz mappings. Hence, these latter are always reached along piecewise constant curves of indices  $t \in \mathbb{R}_+ \mapsto i(t) \in \{1, \dots, N\}$ , which greatly simplifies the analysis of differential estimates involving various geometric quantities of the system. In the graph limit framework, however, these suprema are a priori not attained, and while it may be possible to build piecewise constant curves of indices  $t \in \mathbb{R}_+ \mapsto i(t) \in I$  whose images lie arbitrarily close to the desired quantity, the underlying construction would rely on careful adaptations of measurable selection theorems (see e.g. [3, Theorem 8.3.1]) to the preimages of  $\varepsilon$ -tubes built around curves of suprema.*

*In this context, the application of Scorza-Dragoni's theorem allows to consider well-defined maxima of continuous functions defined over compact sets instead of essential suprema. The time-derivatives of these latter can then be estimated in a fashion similar to that of classical finite-dimensional approaches, up to using Danskin's theorem for the differentiation part, and there only remains to properly account for the error terms stemming from the inner approximation while passing to the limit as  $\varepsilon \rightarrow 0^+$ .*

### 3 Consensus formation via diameter estimates

In this section, we study sufficient conditions for the exponential decay of the diameter of a graphon model. We start in Section 3.1 by a preliminary discussion on scrambling coefficients, and study the formation of  $L^\infty$ -consensus in Section 3.2.

### 3.1 Generalised scrambling coefficient for graphon models

As already discussed in the Introduction, the quantity that usually dictates the asymptotic behaviour of the diameter of a multi-agent system is the so-called *scrambling coefficient*. We point the reference monograph [71] for a detailed overview of its properties.

**Scrambling coefficients in finite and infinite dimension.** Given an adjacency matrix  $\mathbf{A}_N := (a_{ij})_{1 \leq i, j \leq N} \in [0, 1]^{N \times N}$  encoding the interactions of a discrete system of  $N \geq 1$  agents, the scrambling coefficient is defined by

$$\eta(\mathbf{A}_N) := \min_{1 \leq i, j \leq N} \frac{1}{N} \sum_{k=1}^N \min\{a_{ik}, a_{jk}\}. \quad (3.1)$$

Since the self-interactions of the agents are set to be  $a_{ii} = 1$  for each  $i \in \{1, \dots, N\}$ , the latter can be equivalently rewritten as

$$\eta(\mathbf{A}_N) = \min_{1 \leq i, j \leq N} \frac{1}{N} \left( \sum_{k=1, k \neq i, j}^N \min\{a_{ik}, a_{jk}\} + a_{ij} + a_{ji} \right). \quad (3.2)$$

What this formula expresses is that the scrambling coefficient of an interaction topology is positive if and only if, for every pair of agents  $i, j \in \{1, \dots, N\}$ , either  $i$  and  $j$  are communicating with one another, or there exists a third party agent  $k \in \{1, \dots, N\} \setminus \{i, j\}$  which is followed by both  $i$  and  $j$ . We provide in Figure 3.1 three simple examples of interaction topologies with positive scrambling.

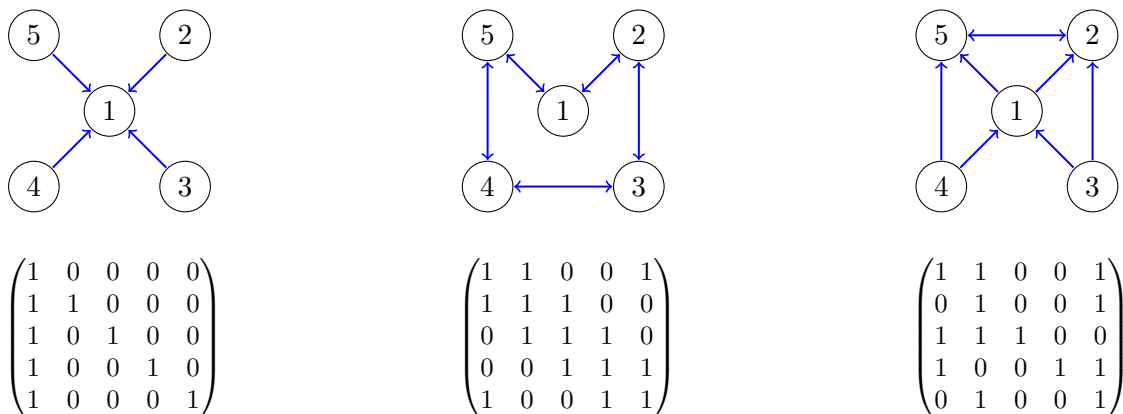


Figure 3.1: Three examples of interaction topologies with positive scrambling for graphs with  $N = 5$  vertices. The interaction graphs are represented on top, and the corresponding adjacency matrices right below.

Given an adjacency operator  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  with kernel  $a \in L^\infty(I \times I, [0, 1])$ , we introduce the following infinite-dimensional generalisation of the scrambling coefficient

$$\eta(\mathcal{A}) := \inf_{i, j \in I} \int_I \min\{a(i, k), a(j, k)\} dk. \quad (3.3)$$

As we shall see in Theorem 3.3 below, this macroscopic quantity does indeed quantify the decay of the diameter of a graphon dynamics of the form (2.8). It is also worth noting that if  $\mathcal{A}_N \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is the adjacency operator associated with a piecewise constant kernel  $a^N \in L^\infty(I \times I, [0, 1])$  of the form (2.6) for some  $N \geq 1$ , then  $\eta(\mathcal{A}_N) = \eta(\mathbf{A}_N)$  as in (3.1), where  $\mathbf{A}_N \in \mathbb{R}^{N \times N}$  is the adjacency matrix of the underlying digraph. This shows that the definition proposed in (3.3) is indeed a proper generalisation of the usual notion of scrambling coefficient.

**The role of stochasticity in the literature.** In most of the existing works studying the asymptotic properties of the diameter in multi-agent dynamics (see e.g. [54, 60, 61] and references therein),

the proof of the decay estimates explicitly rely on the assumption that the adjacency matrices are *stochastic*, namely that they satisfy

$$\frac{1}{N} \sum_{j=1}^N a_{ij} = 1,$$

for each  $i \in \{1, \dots, N\}$ . Indeed, observing that a multi-agent dynamics of the form

$$\dot{x}_i(t) = \frac{1}{N} \sum_{j=1}^N a_{ij}(x_j(t) - x_i(t)),$$

is independent of the values of the diagonal coefficients  $(a_{11}, \dots, a_{NN})$ , one can equivalently consider that any of its solutions is driven by the stochastic interaction weights  $(\tilde{a}_{ij})_{1 \leq i, j \leq N}$ , defined by

$$\tilde{a}_{ii} := N - \sum_{k=1, k \neq i}^N a_{ik} \quad \text{and} \quad \tilde{a}_{ij} := a_{ij}, \quad (3.4)$$

for every  $i, j \in \{1, \dots, N\}$ . In the following proposition, we show that the scrambling coefficient of an interaction topology does not change when the communication weights are redefined using (3.4). While this result is not difficult to establish, it is crucial in order to ensure that the latter transformation does not modify important qualitative properties of the dynamics.

**Proposition 3.1** (Invariance of scrambling coefficients under stochastic reparametrisation). *Let  $\mathbf{A}_N := (a_{ij})_{1 \leq i, j \leq N} \in [0, 1]^{N \times N}$  be the adjacency matrix of an interaction graph, and  $\tilde{\mathbf{A}}_N := (\tilde{a}_{ij})_{1 \leq i, j \leq N}$  be its stochastic reparametrisation obtained via the transformation (3.4). Then,  $\eta(\tilde{\mathbf{A}}_N) = \eta(\mathbf{A}_N)$ .*

*Proof.* Observe first that since  $(a_{ij})_{1 \leq i, j \leq N} \in [0, 1]^{N \times N}$ , it necessarily holds

$$a_{ij} \leq 1 \leq N - \sum_{k=1, k \neq i}^N a_{ik} \quad \text{and} \quad a_{ji} \leq 1 \leq N - \sum_{k=1, k \neq j}^N a_{jk},$$

for every  $i, j \in \{1, \dots, N\}$ . Thus by definition (3.4) of the coefficients of the modified adjacency matrix  $\tilde{\mathbf{A}}_N$ , one has that

$$\begin{aligned} \eta(\tilde{\mathbf{A}}_N) &= \min_{1 \leq i, j \leq N} \frac{1}{N} \left( \sum_{k=1, k \neq i, j}^N \min\{a_{ik}, a_{jk}\} + \min \left\{ a_{ij}, N - \sum_{k=1, k \neq i}^N a_{ik} \right\} \right. \\ &\quad \left. + \min \left\{ N - \sum_{k=1, k \neq j}^N a_{jk}, a_{ji} \right\} \right) \\ &= \min_{1 \leq i, j \leq N} \frac{1}{N} \left( \sum_{k=1, k \neq i, j}^N \min\{a_{ik}, a_{jk}\} + a_{ij} + a_{ji} \right) \\ &= \eta(\mathbf{A}_N), \end{aligned}$$

where we used the alternative expression (3.2) of the scrambling coefficient.  $\square$

**Remark 3.2** (Obstructions to stochastic reparametrisation for graphons). *While the approach described above is very useful in order to simplify computations when considering finite-dimensional multi-agent systems, it is not readily applicable to graphon dynamics. Indeed, one can easily check that the transformation of the adjacency matrix displayed in (3.4) does not admit a well-defined point-wise limit as  $N \rightarrow +\infty$ . On the other hand, it can be shown that the corresponding piecewise constant interaction maps defined as in (2.6) would converge in the sense of distributions towards the sum of a kernel  $a \in L^\infty(I \times I, [0, 1])$  and a Dirac measure supported on the diagonal. Therefore, the resulting right-hand side falls outside the scope of classical graphon dynamics, in which the adjacency operator is simply given by the integral of an  $\mathcal{L}^1_I$ -measurable kernel.*



### 3.2 Exponential consensus formation for topologies with persistent scramblings

In Theorem 3.3 below, we derive a general contraction estimate on the diameter of graphon dynamics, expressed in terms of the generalised scrambling coefficient. In what follows for  $\mathcal{L}^1$ -almost every  $t \geq 0$ , we denote by  $\mathcal{A}(t) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  the adjacency operator with kernel  $a(t) \in L^\infty(I \times I, [0, 1])$ .

**Theorem 3.3** (A general contraction estimate for the diameter). *Fix a radius  $R > 0$  and an initial datum  $x^0 \in L^\infty(I, \mathbb{R}^d)$  satisfying  $\|x^0\|_{L^\infty(I, \mathbb{R}^d)} \leq R$ . Moreover, assume that hypotheses **(GD)** hold, and consider a solution  $x(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, L^2(I, \mathbb{R}^d))$  of (2.7).*

*Then for all times  $t \geq 0$ , it holds that*

$$\mathcal{D}(t) \leq \mathcal{D}(0) \exp\left(-\gamma_R \int_0^t \eta(\mathcal{A}(s)) ds\right),$$

with  $\gamma_R := \min_{r \in [0, 2R]} \phi(r)$ , and where the scrambling coefficient  $\eta(\mathcal{A}(s))$  generated by the communication weights is defined as in (3.3) for  $\mathcal{L}^1$ -almost every  $s \geq 0$ .

Before proving Theorem 3.3, we establish a useful geometric lemma which allows for the estimation of the time-derivative of the diameter without requiring the stochasticity of the adjacency operator.

**Lemma 3.4** (A scalar product inequality). *Let  $J \subset I$  be a closed set,  $x \in C^0(J, \mathbb{R}^d)$  and  $i, j \in J$  be a pair of indices such that  $\max_{k, l \in J} |x(k) - x(l)| = |x(i) - x(j)|$ . Then, one has that*

$$\begin{aligned} \langle x(j), x(i) - x(j) \rangle &= \min_{k \in J} \langle x(k), x(i) - x(j) \rangle \\ &\leq \max_{k \in J} \langle x(k), x(i) - x(j) \rangle = \langle x(i), x(i) - x(j) \rangle. \end{aligned} \tag{3.5}$$

*Proof.* It is a classical result in geometric analysis that the diameter of any subset of  $\mathbb{R}^d$  coincides with the diameter of its convex hull. Thus, defining  $x_\lambda(k) := (1 - \lambda)x(i) + \lambda x(k)$  for some  $k \in J$  and every  $\lambda \in [0, 1]$ , it necessarily holds

$$|x_\lambda(k) - x(j)|^2 \leq |x(i) - x(j)|^2, \tag{3.6}$$

by construction of the indices  $i, j \in J$ . An easy computation then shows that

$$\begin{aligned} |x_\lambda(k) - x(j)|^2 - |x(i) - x(j)|^2 &= \left\langle (x_\lambda(k) - x(j)) - (x(i) - x(j)), (x_\lambda(k) - x(j)) + (x(i) - x(j)) \right\rangle \\ &= \lambda^2 |x(i) - x(k)|^2 + 2\lambda \langle x(k) - x(i), x(i) - x(j) \rangle, \end{aligned}$$

so that (3.6) can be equivalently reformulated as

$$\lambda^2 |x(i) - x(k)|^2 + 2\lambda \langle x(k) - x(i), x(i) - x(j) \rangle \leq 0, \tag{3.7}$$

for every  $\lambda \in [0, 1]$ . Dividing by  $\lambda \in (0, 1]$  and letting  $\lambda \rightarrow 0^+$  in (3.7), we therefore obtain

$$\langle x(k), x(i) - x(j) \rangle \leq \langle x(i), x(i) - x(j) \rangle,$$

for every  $k \in J$ , which is equivalent to stating that  $\langle x(i), x(i) - x(j) \rangle = \max_{k \in J} \langle x(k), x(i) - x(j) \rangle$ . The other inequality in (3.5) can be proven by repeating the same arguments.  $\square$

*Proof of Theorem 3.3.* Let  $x \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, L^2(I, \mathbb{R}^d))$  be a solution of (2.7), and fix a parameter  $\varepsilon > 0$ . Then by Scorza-Dragoni's theorem (see Theorem 2.1 above), there exists a compact set  $I_\varepsilon \subset I$  with  $\mathcal{L}_I^1(I \setminus I_\varepsilon) < \varepsilon$  such that the restriction  $x : \mathbb{R}_+ \times I_\varepsilon \rightarrow \mathbb{R}^d$  is a continuous map. We then consider the restricted diameters, defined by

$$\mathcal{D}_\varepsilon(t) := \max_{i, j \in I_\varepsilon} |x(t, i) - x(t, j)|$$

for all times  $t \geq 0$ , and introduce the nonempty subset of pairs of indices  $\Pi_\varepsilon(t) \subset I_\varepsilon \times I_\varepsilon$  for which this maximum is reached.

Observe first that  $\mathcal{D}_\varepsilon(\cdot)$  is defined as the pointwise maximum of a family of equi-locally Lipschitz mappings, and is therefore a locally Lipschitz function. Hence by Rademacher's theorem (see e.g. [33, Theorem 3.2]), it is differentiable  $\mathcal{L}^1$ -almost everywhere, and it holds by Danskin's theorem (see Theorem 2.2 above) that

$$\frac{1}{2} \frac{d}{dt} \mathcal{D}_\varepsilon(t)^2 = \max_{(i,j) \in \Pi_\varepsilon(t)} \langle \partial_t x(t, i) - \partial_t x(t, j), x(t, i) - x(t, j) \rangle. \quad (3.8)$$

We now fix a pair of indices  $(i, j) \in \Pi_\varepsilon(t)$ , and aim at evaluating the right-hand side of (3.8). Recalling that  $x(\cdot)$  solves (2.7), one has that

$$\begin{aligned} & \langle \partial_t x(t, i) - \partial_t x(t, j), x(t, i) - x(t, j) \rangle \\ &= \int_I a(t, i, k) \phi(|x(t, i) - x(t, k)|) \langle x(t, k) - x(t, i), x(t, i) - x(t, j) \rangle dk \\ & \quad - \int_I a(t, j, k) \phi(|x(t, j) - x(t, k)|) \langle x(t, k) - x(t, j), x(t, i) - x(t, j) \rangle dk \\ &= \int_{I_\varepsilon} a(t, i, k) \phi(|x(t, i) - x(t, k)|) \langle x(t, k) - x(t, i), x(t, i) - x(t, j) \rangle dk \\ & \quad - \int_{I_\varepsilon} a(t, j, k) \phi(|x(t, j) - x(t, k)|) \langle x(t, k) - x(t, j), x(t, i) - x(t, j) \rangle dk + \varepsilon C(t), \end{aligned} \quad (3.9)$$

where  $C(\cdot) \in L_{\text{loc}}^\infty(\mathbb{R}_+, \mathbb{R})$ . Upon defining the quantity

$$\eta_{ij}(t, k) := \min \left\{ a(t, i, k) \phi(|x(t, i) - x(t, k)|), a(t, j, k) \phi(|x(t, j) - x(t, k)|) \right\},$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$  and  $\mathcal{L}^1_I$ -almost every  $k \in I$ , one can rewrite (3.9) as

$$\begin{aligned} & \langle \partial_t x(t, i) - \partial_t x(t, j), x(t, i) - x(t, j) \rangle \\ &= \int_{I_\varepsilon} a(t, i, k) \phi(|x(t, i) - x(t, k)|) \langle x(t, k), x(t, i) - x(t, j) \rangle dk \\ & \quad - \int_{I_\varepsilon} a(t, j, k) \phi(|x(t, j) - x(t, k)|) \langle x(t, k), x(t, i) - x(t, j) \rangle dk \\ & \quad - \left( \int_{I_\varepsilon} a(t, i, k) \phi(|x(t, i) - x(t, k)|) dk \right) \langle x(t, i), x(t, i) - x(t, j) \rangle \\ & \quad + \left( \int_{I_\varepsilon} a(t, j, k) \phi(|x(t, j) - x(t, k)|) dk \right) \langle x(t, j), x(t, i) - x(t, j) \rangle + \varepsilon C(t) \\ &= \int_{I_\varepsilon} \left( a(t, i, k) \phi(|x(t, i) - x(t, k)|) - \eta_{ij}(t, k) \right) \langle x(t, k), x(t, i) - x(t, j) \rangle dk \\ & \quad + \int_{I_\varepsilon} \left( \eta_{ij}(t, k) - a(t, j, k) \phi(|x(t, j) - x(t, k)|) \right) \langle x(t, k), x(t, i) - x(t, j) \rangle dk \\ & \quad - \left( \int_{I_\varepsilon} a(t, i, k) \phi(|x(t, i) - x(t, k)|) dk \right) \langle x(t, i), x(t, i) - x(t, j) \rangle \\ & \quad + \left( \int_{I_\varepsilon} a(t, j, k) \phi(|x(t, j) - x(t, k)|) dk \right) \langle x(t, j), x(t, i) - x(t, j) \rangle + \varepsilon C(t), \end{aligned} \quad (3.10)$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ . Observe now that by construction of the parameters  $\eta_{ij}(t, k)$ , the first two

terms in the right-hand side of (3.10) can be estimated as follows

$$\begin{aligned}
& \int_{I_\varepsilon} \left( a(t, i, k) \phi(|x(t, i) - x(t, k)|) - \eta_{ij}(t, k) \right) \langle x(t, k), x(t, i) - x(t, j) \rangle dk \\
& \quad + \int_I \left( \eta_{ij}(t, k) - a(t, j, k) \phi(|x(t, j) - x(t, k)|) \right) \langle x(t, k), x(t, i) - x(t, j) \rangle dk \\
& \leq \left( \int_{I_\varepsilon} \left( a(t, i, k) \phi(|x(t, i) - x(t, k)|) - \eta_{ij}(t, k) \right) dk \right) \max_{k \in I_\varepsilon} \langle x(t, k), x(t, i) - x(t, j) \rangle \\
& \quad - \left( \int_{I_\varepsilon} \left( a(t, j, k) \phi(|x(t, j) - x(t, k)|) - \eta_{ij}(t, k) \right) dk \right) \min_{k \in I_\varepsilon} \langle x(t, k), x(t, i) - x(t, j) \rangle \\
& \leq \left( - \int_{I_\varepsilon} \eta_{ij}(t, k) dk \right) \max_{k, l \in I_\varepsilon} \langle x(t, k) - x(t, l), x(t, i) - x(t, j) \rangle \\
& \quad + \left( \int_{I_\varepsilon} a(t, i, k) \phi(|x(t, i) - x(t, k)|) dk \right) \max_{k \in I_\varepsilon} \langle x(t, k), x(t, i) - x(t, j) \rangle \\
& \quad - \left( \int_{I_\varepsilon} a(t, j, k) \phi(|x(t, j) - x(t, k)|) dk \right) \min_{k \in I_\varepsilon} \langle x(t, k), x(t, i) - x(t, j) \rangle \\
& \leq \left( - \inf_{i, j \in I} \int_I \eta_{ij}(t, k) dk + c_\phi \varepsilon \right) \mathcal{D}_\varepsilon(t)^2 \\
& \quad + \left( \int_{I_\varepsilon} a(t, i, k) \phi(|x(t, i) - x(t, k)|) dk \right) \max_{k \in I_\varepsilon} \langle x(t, k), x(t, i) - x(t, j) \rangle \\
& \quad - \left( \int_{I_\varepsilon} a(t, j, k) \phi(|x(t, j) - x(t, k)|) dk \right) \min_{k \in I_\varepsilon} \langle x(t, k), x(t, i) - x(t, j) \rangle,
\end{aligned} \tag{3.11}$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ , where we recall that  $c_\phi = \sup_{r \in \mathbb{R}_+} \phi(r) < +\infty$  is finite as a consequence of **(GD)**-(i). Remark now that by Lemma 3.4, it further holds

$$\begin{aligned}
\langle x(t, j), x(t, i) - x(t, j) \rangle &= \min_{k \in I_\varepsilon} \langle x(t, k), x(t, i) - x(t, j) \rangle \\
&\leq \max_{k \in I_\varepsilon} \langle x(t, k), x(t, i) - x(t, j) \rangle = \langle x(t, i), x(t, i) - x(t, j) \rangle,
\end{aligned} \tag{3.12}$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$  and  $\mathcal{L}^1_{\mathbb{L}}$ -almost every  $k \in I_\varepsilon$ , where we used the fact that  $(i, j) \in \Pi_\varepsilon(t)$ . Hence, by plugging (3.10), (3.11) and (3.12) into (3.8) while observing that

$$\eta_{ij}(t, k) \geq \gamma_R \min \{ a(t, i, k), a(t, j, k) \},$$

and also that the former series of estimates hold true for every pair of indices  $(i, j) \in \Pi_\varepsilon(t)$ , we recover the differential inequality

$$\frac{1}{2} \frac{d}{dt} \mathcal{D}_\varepsilon(t)^2 \leq \left( -\gamma_R \eta(\mathcal{A}(t)) + c_\phi \varepsilon \right) \mathcal{D}_\varepsilon(t)^2 + \varepsilon C(t),$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ . By applying the differential version of Grönwall's lemma, we then deduce

$$\begin{aligned}
\mathcal{D}_\varepsilon(t)^2 &\leq \mathcal{D}_\varepsilon(0)^2 \exp \left( -2\gamma_R \int_0^t \eta(\mathcal{A}(s)) ds + 2c_\phi \varepsilon t \right) \\
&\quad + 2\varepsilon \int_0^t C(s) \exp \left( -2\gamma_R \int_s^t \eta(\mathcal{A}(\sigma)) d\sigma - 2c_\phi \varepsilon (t - s) \right) ds,
\end{aligned} \tag{3.13}$$

for all times  $t \geq 0$ . Recall now that by Lemma 2.3, we have the pointwise convergence

$$\mathcal{D}_\varepsilon(t) \xrightarrow{\varepsilon \rightarrow 0^+} \mathcal{D}(t),$$

and thus by passing to the limit as  $\varepsilon \rightarrow 0^+$  in (3.13), we recover the desired exponential decay estimate

$$\mathcal{D}(t) \leq \mathcal{D}(0) \exp \left( -\gamma_R \int_0^t \eta(\mathcal{A}(s)) ds \right),$$

for all times  $t \geq 0$ , which concludes the proof of Theorem 3.3.  $\square$

**Remark 3.5** (Comparing our approach with previous diameter estimates in the literature). *As mentioned above, the existing proofs of diameter contractions for finite-dimensional multi-agent systems involving scrambling coefficient (see e.g. [60, 61]) presuppose that the adjacency matrices  $\mathbf{A}_N(t)$  of the system are stochastic for all times  $t \geq 0$ . In practice, this simplifying assumption allows to write meaningful differential estimates on the quantity  $t \in \mathbb{R}_+ \mapsto |x_i(t) - x_j(t)| \in \mathbb{R}_+$  for every pair of indices  $i, j \in \{1, \dots, N\}$ . Hereinabove on the contrary, we only estimate the derivative of this quantity for pairs of indices which realise the (relaxed) diameter, and use the geometric information of Lemma 3.4 to take care of the terms that were usually simplified using the stochasticity assumption. In this sense, our proof strategy provides an alternative and perhaps more general diameter estimate for multi-agent systems, as it is more natural and valid for both finite- and infinite-dimensional models.*

Based on this general diameter estimate, it is possible to establish the formation of  $L^\infty$ -consensus in the case where the scrambling coefficients are persistent in a suitable sense.

**Theorem 3.6** (Exponential  $L^\infty$ -consensus formation under persistent scrambling). *Assume that the hypotheses of Theorem 3.3 are satisfied, and suppose that there exists a pair of coefficients  $(\tau, \mu) \in \mathbb{R}_+^* \times (0, 1]$  such that the persistence condition*

$$\frac{1}{\tau} \int_t^{t+\tau} \eta(\mathcal{A}(s)) ds \geq \mu, \quad (3.14)$$

holds for all times  $t \geq 0$ .

Then, there exists an element  $x^\infty \in \mathcal{C}(x^0)$  and constants  $\alpha, \gamma > 0$  depending only on  $(\phi(\cdot), R, \tau, \mu)$  such that

$$\|x(t) - x^\infty\|_{L^\infty(I, \mathbb{R}^d)} \leq \alpha \mathcal{D}(0) \exp(-\gamma \mu t), \quad (3.15)$$

for all times  $t \geq 0$ . In particular, the solutions of (2.7) exponentially converge to consensus in the  $L^\infty(I, \mathbb{R}^d)$ -norm topology.

*Proof.* Start by fixing an element  $i \in I$ , and observe that for every  $0 \leq t_1 \leq t_2 < +\infty$ , one has that

$$\begin{aligned} |x(t_2, i) - x(t_1, i)| &\leq \int_{t_1}^{t_2} \int_I a(s, i, j) \phi(|x(s, i) - x(s, j)|) |x(s, j) - x(s, i)| dj ds \\ &\leq c_\phi \int_{t_1}^{t_2} \mathcal{D}(s) ds \\ &\leq c_\phi \int_{t_1}^{+\infty} \mathcal{D}(0) \exp\left(-\gamma_R \int_0^s \eta(\mathcal{A}(\sigma)) d\sigma\right) ds, \end{aligned} \quad (3.16)$$

where we recall that  $c_\phi = \sup_{r \in \mathbb{R}_+} \phi(r) < +\infty$ . Then, as a consequence of the persistence condition (3.14), it further holds that

$$\begin{aligned} \int_0^s \eta(\mathcal{A}(\sigma)) d\sigma &\geq \int_0^{\lfloor s/\tau \rfloor \tau} \eta(\mathcal{A}(\sigma)) d\sigma \\ &\geq \sum_{k=0}^{\lfloor s/\tau \rfloor - 1} \int_{k\tau}^{(k+1)\tau} \eta(\mathcal{A}(\sigma)) d\sigma \geq \mu(s - \tau), \end{aligned} \quad (3.17)$$

for all times  $s \in [t_1, t_2]$ , where  $\lfloor \bullet \rfloor$  denotes the lower integer part of a real number. Thus, by plugging (3.17) into (3.16), we obtain

$$|x(t_2, i) - x(t_1, i)| \leq c_\phi \int_{t_1}^{+\infty} \mathcal{D}(0) \exp\left(-\gamma_R \mu(s - \tau)\right) ds \xrightarrow{t_1, t_2 \rightarrow +\infty} 0, \quad (3.18)$$

which yields the existence of a map  $x^\infty \in L^\infty(I, \mathbb{R}^d)$  such that  $x(t, i) \rightarrow x^\infty(i)$  as  $t \rightarrow +\infty$  for  $\mathcal{L}^1_I$ -almost every  $i \in I$ . Upon noticing that the convergence in (3.18) is uniform with respect to  $i \in I$ , we further have

$$\|x(t) - x^\infty\|_{L^\infty(I, \mathbb{R}^d)} \xrightarrow{t \rightarrow +\infty} 0. \quad (3.19)$$

We now choose any pair of indices  $i, j \in I$ , and remark that for all times  $t \geq 0$ , it holds

$$|x^\infty(i) - x^\infty(j)| \leq |x^\infty(i) - x(t, i)| + |x(t, j) - x^\infty(j)| + \mathcal{D}(t).$$

By combining the diameter contraction of Theorem 3.3 together with the convergence results of (3.18) and (3.19), we obtain that  $x^\infty(i) = x^\infty(j)$  for  $\mathcal{L}_{\perp I}^1$ -almost every  $i, j \in I$ . Therefore, the map  $x^\infty$  is in fact constant, and by Proposition 2.11 it also belongs to the initial convex hull  $\mathcal{C}(x^0)$  of the system.

Observing that  $x^\infty \in \mathcal{C}(x^0)$ , and using the fact that the diameter of a set coincides with that of its convex hull together with the series of estimates (3.17), we finally get

$$|x(t, i) - x^\infty| \leq \mathcal{D}(t) \leq \mathcal{D}(0) \exp\left(-\gamma_R \mu(t - \tau)\right),$$

for all times  $t \geq 0$  and  $\mathcal{L}_{\perp I}^1$ -almost every  $i \in I$ . Taking the essential supremum over  $I$  in the previous expression then yields the exponential contraction estimate (3.15) with  $\alpha := \exp(\gamma_R \mu \tau)$  and  $\gamma := \gamma_R$ .  $\square$

**Remark 3.7** (On the nature of the asymptotic consensus point). *Observe that, in general, the emerging consensus  $x^\infty \in \mathcal{C}(x^0)$  – whose existence follows from Theorem 3.6 – does not necessarily coincide with the initial barycenter  $\bar{x}^0 := \int_I x^0(i) di$  of the system. This is due to the fact that the latter may not be an intrinsic constant of motion for asymmetric topologies (see also Section 4.2.3 below).*

## 4 Consensus formation via energy estimates

In this section, we investigate another family of sufficient conditions ensuring the asymptotic formation of  $L^2$ -consensus for time-dependent graphon models. We first start in Section 4.1 by a general discussion on the notion of *algebraic connectivity*, along with a generalisation of this classical object to graphon models. We then study in Section 4.2 the exponential convergence to consensus for topologies that are symmetric, balanced, and made of a disjoint union of strongly connected components.

### 4.1 Generalised algebraic connectivity for graphon models

In this section, we start by a discussion on the notion of algebraic connectivity for finite-dimensional collective dynamics, and then propose a natural generalisation for graphon models.

**Graph-Laplacians in the finite-dimensional setting.** It is widely known (see e.g. [59, 64] along with the monographs [6, 57]) that in order to study the formation of consensus with respect to the  $\ell^2$ -norm for discrete systems of the form (2.4), it is often insightful to introduce the *graph-Laplacian operator*  $\mathbf{L}_N : \mathbb{R}_+ \times (\mathbb{R}^d)^N \rightarrow \mathcal{L}((\mathbb{R}^d)^N)$  associated with the interaction topology, defined by

$$(\mathbf{L}_N(t, \mathbf{x})\mathbf{y})_i := \frac{1}{N} \sum_{j=1}^N a_{ij}(t) \phi(|x_i - x_j|) (y_i - y_j) \quad \text{for each } i \in \{1, \dots, N\}, \quad (4.1)$$

for  $\mathcal{L}^1$ -almost every  $t \in \mathbb{R}_+$  and all  $\mathbf{x}, \mathbf{y} \in (\mathbb{R}^d)^N$ . This permits to rewrite (2.4) as the semilinear evolution equation

$$\dot{\mathbf{x}}(t) = -\mathbf{L}_N(t, \mathbf{x}(t))\mathbf{x}(t), \quad \mathbf{x}(0) = \mathbf{x}^0, \quad (4.2)$$

posed in the configuration space  $(\mathbb{R}^d)^N$ , and to study the convergence to consensus in terms of the eigenvalues of the matrices  $\mathbf{L}_N(t, \mathbf{x}(t))$  using techniques of stability theory (see e.g. [63]). A first key observation about graph-Laplacians is that their kernel is never trivial, as it can be checked that  $\mathbf{L}_N(t, \mathbf{x})\mathbf{y} = 0$  whenever  $\mathbf{y} \in (\mathbb{R}^d)^N$  belongs to the so-called *consensus manifold*, that is defined by

$$\mathcal{C}_N := \left\{ \mathbf{x} \in (\mathbb{R}^d)^N \text{ s.t. } x_1 = \dots = x_N \right\}. \quad (4.3)$$

This implies in particular that 0 is always a nontrivial eigenvalue of  $\mathbf{L}_N(t, \mathbf{x})$  with multiplicity at least  $d \geq 1$ . Because the consensus manifold is also the set of equilibrium points of the dynamics (2.4), the convergence of a solution  $\mathbf{x}(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, (\mathbb{R}^d)^N)$  towards consensus can be quantified in terms of the  $(d + 1)$ -th smallest eigenvalue of the matrix  $\mathbf{L}_N(t, \mathbf{x}(t))$ , counted with multiplicity.

**Various notions of algebraic connectivity for discrete systems.** The notion of *algebraic connectivity* was originally introduced in the seminal work [34] for symmetric adjacency matrices  $\mathbf{A}_N := (a_{ij})_{1 \leq i, j \leq N} \in [0, 1]^{N \times N}$  – namely when  $a_{ij} = a_{ji}$  for every pair of indices  $i, j \in \{1, \dots, N\}$  –, and defined as the  $(d + 1)$ -th smallest eigenvalue of the corresponding graph-Laplacian operator  $\mathbf{L}_N \in \mathcal{L}((\mathbb{R}^d)^N)$ , counted with multiplicity. By the Courant-Fischer min-max theorem (see e.g. [43, Theorem 4.2.12]), the latter could be alternatively defined as

$$\lambda_2(\mathbf{L}_N) := \min_{\mathbf{x} \in \mathcal{C}_N^\perp} \frac{\langle \mathbf{L}_N \mathbf{x}, \mathbf{x} \rangle}{|\mathbf{x}|^2} \geq 0, \quad (4.4)$$

where the orthogonal complement of the consensus manifold can be characterised explicitly by

$$\mathcal{C}_N^\perp = \left\{ \mathbf{x} \in (\mathbb{R}^d)^N \text{ s.t. } \bar{\mathbf{x}} = 0 \right\}. \quad (4.5)$$

Here,  $\bar{\mathbf{x}} := \frac{1}{N} \sum_{i=1}^N x_i \in \mathbb{R}^d$  stands for the mean value of an element  $\mathbf{x} \in (\mathbb{R}^d)^N$ . It has since been a classical result in algebraic graph theory that  $\lambda_2(\mathbf{L}_N) > 0$  if and only if the underlying undirected graph is *connected* in the sense of Definition 2.4 above, and that the multiplicity of the zero eigenvalue of  $\mathbf{L}_N$  is equal to the product between the number of connected components in the topology and the dimension  $d$  of the state space (see e.g. the reference paper [58]).

When the interaction topology of the system is asymmetric, it is still possible to adapt the notion of algebraic connectivity in a meaningful way. In [79], it was shown that the definition given in (4.4) still makes sense in the class of *balanced topologies*, which are characterised by the fact that

$$\frac{1}{N} \sum_{j=1}^N a_{ij} = \frac{1}{N} \sum_{j=1}^N a_{ji}$$

for each  $i \in \{1, \dots, N\}$ . Indeed, it directly follows in this context that  $\mathbf{L}_N^\top \mathbf{x} = 0$  for each  $\mathbf{x} \in \mathcal{C}_N$ , and that the algebraic connectivity of the interaction topology is positive if and only if the interaction graph is strongly connected (see e.g. [79, Lemma 17]). We point the reader to Figure 4.1 for some examples of balanced topologies.

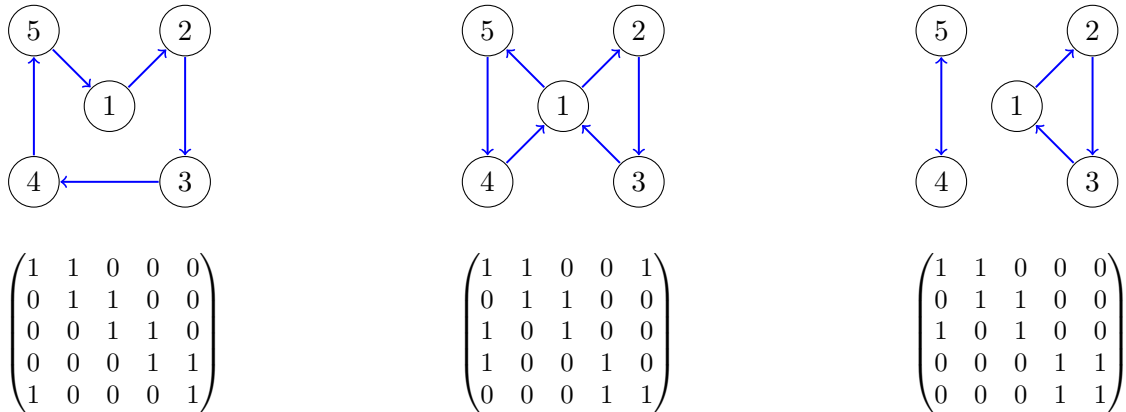


Figure 4.1: Three examples of balanced topologies for graphs with  $N = 5$  vertices. The graphs on the left and in the center are strongly connected, while that on the right is a disjoint union of two strongly connected components. The corresponding adjacency matrices are represented right below each graph.

In the subsequent work [81], it was finally shown that the notion of algebraic connectivity could be further generalised to interaction topologies that can be written as *disjoint unions of strongly connected components*, in the sense of Definition 2.4 above. In this context, the definition of algebraic connectivity relies on the following structural result of Perron-Frobenius theory.

**Theorem 4.1** (Algebraic characterisation of strong connectivity). *Let  $\mathbf{A}_N := (a_{ij})_{1 \leq i, j \leq N} \in [0, 1]^{N \times N}$  be an adjacency matrix and  $\mathbf{L}_N \in \mathcal{L}((\mathbb{R}^d)^N)$  be the corresponding graph-Laplacian operator. Then, the*

underlying interaction graph is a disjoint union of strongly connected components if and only if there exists an element  $\mathbf{v} := (v_i)_{1 \leq i \leq N} \in (\mathbb{R}_+^*)^N$  such that

$$\mathbf{L}_N^\top \mathbf{v} = 0 \quad \text{and} \quad \frac{1}{N} \sum_{i=1}^N v_i = 1, \quad (4.6)$$

where  $\mathbf{v} := (v_1, \dots, v_1, \dots, v_N, \dots, v_N) \in (\mathbb{R}^d)^N$ .

*Proof.* The proof of this result for  $d = 1$  can be found e.g. in [81, Lemma 1]. Then for an arbitrary state-space dimensions  $d \geq 1$ , it is sufficient to define  $\mathbf{v} \in (\mathbb{R}^d)^N$  as indicated above.  $\square$

**Remark 4.2** (1-dimensional versus  $d$ -dimensional linear systems). *When the right-hand side of a multi-agent system is linear, there is no coupling between the dynamics of the components  $(x_i^k)_{1 \leq k \leq d}$  of the agents for  $i \in \{1, \dots, N\}$ . This explains why the canonical eigenvector  $\mathbf{v} \in (\mathbb{R}^d)^N$  for the  $d$ -dimensional system is simply obtained in this context by concatenating  $d$  copies of that generated by the action of  $\mathbf{L}_N$  on a 1-dimensional space.*

Assuming that the interaction topology described by the adjacency matrix  $\mathbf{A}_N$  is a disjoint union of strongly connected components, we consider the following weighted version of the graph-Laplacian

$$\mathbf{L}_N^{\mathbf{v}} := \text{diag}(\mathbf{v}) \mathbf{L}_N,$$

where  $\text{diag}(\mathbf{v}) \in \mathcal{L}((\mathbb{R}^d)^N)$  is the diagonal matrix whose entries are the components of  $\mathbf{v} \in (\mathbb{R}^d)^N$ . Then, the algebraic connectivity of the underlying topology is defined by

$$\lambda_2(\mathbf{L}_N^{\mathbf{v}}) = \min_{\mathbf{x} \in \mathcal{C}_N^\perp} \frac{\langle \mathbf{L}_N^{\mathbf{v}} \mathbf{x}, \mathbf{x} \rangle}{|\mathbf{x}|^2} \geq 0, \quad (4.7)$$

and it can be shown that this definition carries with it all the interesting properties of  $\lambda_2(\cdot)$  (see [80, 81]). In particular, it is shown in [81, Section 4] that in this context,  $\lambda_2(\mathbf{L}_N^{\mathbf{v}}) > 0$  if and only if the interaction graph is strongly connected. It is also worth noting that in the particular case of balanced topologies, the canonical vector given by Theorem 4.1 is simply  $\mathbf{v} := (1, \dots, 1) \in (\mathbb{R}^d)^N$ , so that the definitions of (4.4) and (4.7) coincide.

**Remark 4.3** (Comparison between the algebraic connectivity and the scrambling coefficient). *Given a symmetric adjacency matrix  $\mathbf{A}_N \in [0, 1]^{N \times N}$ , it is a known fact that  $\lambda_2(\mathbf{L}_N) \geq \eta(\mathbf{A}_N)$ , namely that the scrambling coefficient provides a lower-bound on the algebraic connectivity of the corresponding interaction topology (see e.g. [61, Section 2.2, Example 1]). This, however, is no longer true when the adjacency matrix is asymmetric. Indeed, the balanced interaction topologies displayed at the left and center positions in Figure 4.1 are such that  $\eta(\mathbf{A}_N) = 0$ , as in both situations the vertices  $\{2, 5\}$  are neither connected directly to one another nor to a common third party vertex, yet it can be checked fairly easily that  $\lambda_2(\mathbf{L}_N) > 0$  for both of these graphs. Conversely, the interaction topology exposed at the left of Figure 3.1 is such that  $\eta(\mathbf{A}_N) > 0$ , but it does not admit a well-defined algebraic connectivity.*

**Infinite-dimensional graph-Laplacians.** Similarly to what was done above for finite-dimensional cooperative systems, the graphon dynamics (2.8) can be reformulated as an ordinary differential equation in the configuration space  $L^2(I, \mathbb{R}^d)$ . To this end, we introduce the *graph-Laplacian operator*  $\mathbb{L} : \mathbb{R}_+ \times L^2(I, \mathbb{R}^d) \rightarrow \mathcal{L}(L^2(I, \mathbb{R}^d))$ , defined by

$$(\mathbb{L}(t, x)y)(i) := \int_I a(t, i, j) \phi(|x(i) - x(j)|) (y(i) - y(j)) dj \quad \text{for } \mathcal{L}_I^1\text{-almost every } i \in I, \quad (4.8)$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$  and every  $x, y \in L^2(I, \mathbb{R}^d)$ . In keeping with the notions introduced in Section 2.2, observe that  $\mathbb{L}(t, x)$  can be recast as the difference between the multiplication operator  $\mathcal{M}_d(t, x) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  by the so-called *in-degree function*  $d(t, x) \in L^\infty(I, [0, 1])$ , defined by

$$(\mathcal{M}_d(t, x)y)(i) := \left( \int_I a(t, i, j) \phi(|x(i) - x(j)|) dj \right) y(i) \quad \text{for } \mathcal{L}_I^1\text{-almost every } i \in I,$$

to which one subtracts a kernel-type adjacency operator  $\mathcal{A}(t, x) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  that writes

$$(\mathcal{A}(t, x)y)(i) := \int_I a(t, i, j)\phi(|x(i) - x(j)|)x(j)dj \quad \text{for } \mathcal{L}_{\sqcup I}^1\text{-almost every } i \in I,$$

where both expressions are understood for  $\mathcal{L}^1$ -almost every  $t \geq 0$  and each  $x, y \in L^2(I, \mathbb{R}^d)$ . This allows us to rewrite the graphon counterpart of the dynamics (4.2) as the semilinear Cauchy problem

$$\dot{x}(t) = -\mathbb{L}(t, x(t))x(t), \quad x(0) = x^0, \quad (4.9)$$

formulated in the Hilbert space  $L^2(I, \mathbb{R}^d)$ . In the sequel, we will sometimes consider the partial graph-Laplacian operators  $\mathbb{L}^a(\cdot)$  generated solely by the kernel  $a \in L^\infty(\mathbb{R}_+ \times I \times I, [0, 1])$ , defined by

$$(\mathbb{L}^a(t)y)(i) := \int_I a(t, i, j)(y(i) - y(j))dj \quad \text{for } \mathcal{L}_{\sqcup I}^1\text{-almost every } i \in I \quad (4.10)$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$  and each  $y \in L^2(I, \mathbb{R}^d)$ .

**Generalised algebraic connectivity for graphon models.** Just as finite-dimensional graph-Laplacians always admit a nontrivial kernel, their infinite-dimensional counterparts introduced in (4.8) also vanish on the consensus manifold, which in the context of graphon dynamics is defined by

$$\mathcal{C} := \left\{ x \in L^2(I, \mathbb{R}^d) \text{ s.t. } x \text{ is constant over } I \right\}. \quad (4.11)$$

In other words, the kernel of a graph-Laplacian operator is a nontrivial subspace of dimension at least  $d \geq 1$ . Similarly to the finite-dimensional situation, the configuration space  $L^2(I, \mathbb{R}^d)$  admits a canonical decomposition as the orthogonal sum  $\mathcal{C} \oplus \mathcal{C}^\perp$ . In this context, one also has the characterisation

$$\mathcal{C}^\perp = \left\{ x \in L^2(I, \mathbb{R}^d) \text{ s.t. } \bar{x} = 0 \right\}, \quad (4.12)$$

where  $\bar{x} := \int_I x(i)di \in \mathbb{R}^d$  denotes the *barycenter* of an element  $x \in L^2(I, \mathbb{R}^d)$ . We start our discussion on algebraic connectivities for graphon models by stating in the following definition a natural generalisation of the notion of balanced topology.

**Definition 4.4** (Balanced interaction topologies). *The interaction topology described by an adjacency operator  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is said to be balanced provided that*

$$\int_I a(i, j)di = \int_I a(j, i)di,$$

for  $\mathcal{L}_{\sqcup I}^1$ -almost every  $i \in I$ , or equivalently if the corresponding graph-Laplacian  $\mathbb{L} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  satisfies  $\mathbb{L}^*x = 0$  for each  $x \in \mathcal{C}$ .

Given a general adjacency operator  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  with kernel  $a \in L^\infty(I \times I, [0, 1])$ , we define the corresponding in-degree function as

$$d(i) := \int_I a(i, j)dj, \quad (4.13)$$

for  $\mathcal{L}_{\sqcup I}^1$ -almost every  $i \in I$ , and denote by  $\mathcal{M}_d \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  the multiplication operator by  $d \in L^\infty(I, [0, 1])$ . Then, the graph-Laplacian of the topology simply writes as

$$\mathbb{L} := \mathcal{M}_d - \mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d)). \quad (4.14)$$

In this context, we define the algebraic connectivity of a balanced topology by

$$\lambda_2(\mathbb{L}) := \inf_{x \in \mathcal{C}^\perp} \frac{\langle \mathbb{L}x, x \rangle_{L^2(I, \mathbb{R}^d)}}{\|x\|_{L^2(I, \mathbb{R}^d)}^2} \geq 0, \quad (4.15)$$



and we shall see that it plays for  $(-\mathbb{L})$  a role that is similar to that of the usual *spectral bound* for semigroup generators (see e.g. [32, Chapter IV, Definition 2.1]). Just as for scrambling coefficients, the notion of algebraic connectivity that we propose is stable under discretisation. Indeed, if  $\mathbb{L}_N \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is the graph-Laplacian operator associated with a piecewise constant kernel  $a^N \in L^\infty(I \times I, [0, 1])$  of the form (2.6), then it will hold that  $\lambda_2(\mathbb{L}_N) = \lambda_2(\mathbf{L}_N)$ , where  $\mathbf{L}_N \in \mathcal{L}((\mathbb{R}^d)^N)$  denotes the usual graph-Laplacian matrix of the underlying digraph.

Similarly to the situation described for discrete multi-agent systems, it is possible to define a suitable notion of strong connectivity for graphons, which involves the *Lebesgue points* (see e.g. [33, Definition 1.24]) of the interaction kernel. We recall below the definition of this concept, which is borrowed from the recent article [16].

**Definition 4.5** (Strongly connected interaction topologies). *The interaction topology described by an adjacency operator  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is said to be strongly connected if the following conditions hold.*

- (a) *For  $\mathcal{L}_{\sqcup I}^1$ -almost every  $i \in I$  and for each Lebesgue point  $j \in I \setminus \{i\}$  of  $a(i, \cdot) \in L^\infty(I, [0, 1])$ , there exists an integer  $m \geq 1$  and a finite sequence  $(l_k)_{1 \leq k \leq m} \subset I$  satisfying  $i = l_1$ ,  $j = l_m$  and such that  $l_{k+1} \in \text{supp}(a(l_k, \cdot))$  is a Lebesgue point of  $a(l_k, \cdot)$  for each  $k \in \{1, \dots, m-1\}$ .*
- (b) *One has that  $\delta := \inf_{i \in I} \int_I a(i, j) dj > 0$ .*

Analogously, we say that an interaction topology is a disjoint union of strongly connected components if there exists an at most countable family of sets  $\{I_n\}_{n=1}^{+\infty} \subset \mathcal{P}(I)$  such that  $I = \cup_{n \geq 1}^{+\infty} I_n$ , which satisfies

$$\mathcal{L}_{\sqcup I}^1(I_n) > 0, \quad \mathcal{L}_{\sqcup I}^1(I_n \cap I_m) = 0 \quad \text{and} \quad \text{supp}(a(i_n, \cdot)) \subset I_n \quad \text{for } \mathcal{L}_{\sqcup I}^1\text{-almost every } i_n \in I_n,$$

for each  $m, n \geq 1$  with  $m \neq n$ , and on which the restricted operators  $\mathcal{A}_n \in \mathcal{L}(L^2(I_n, \mathbb{R}^d))$  define strongly connected topologies whose in-degree function comply with the quantitative lower-bound

$$\inf_{i_n \in I_n} \frac{1}{\mathcal{L}_{\sqcup I}^1(I_n)} \int_{I_n} a(i_n, j) dj \geq \delta, \quad (4.16)$$

for a given uniform constant  $\delta > 0$ .

**Remark 4.6** (On the lower-bound (4.16) on the in-degree function). *Observe first that the uniform lower-bound (4.16) imposed on each of the strongly connected components  $I_n$  is automatically satisfied when the family of sets  $\{I_n\}_{n=1}^{+\infty} \subset \mathcal{P}(I)$  is finite. When the latter is infinite countable, it transcribes the fact that while the measure of the sets  $I_n$  must eventually vanish, the in-degree function remains lower-bounded in average. A typical example of topology expressed as an infinite union of strongly connected components satisfying this condition is given by the symmetric kernel defined by*

$$a(i, j) := \sum_{n=1}^{+\infty} \mathbb{1}_{I_n}(i) \mathbb{1}_{I_n}(j),$$

for every  $i, j \in I$ , where the family of sets  $\{I_n\}_{n=1}^{+\infty}$  is given by  $I_n := (\frac{1}{n+1}, \frac{1}{n}]$  for each  $n \geq 1$ .

When an interaction topology is decomposable into a disjoint union of strongly connected components, it is yet again possible to define a proper notion of algebraic connectivity, by resorting to the following adaptation of one of the main results from [16]. The latter can be seen as a generalisation of the “only if” part of Theorem 4.1 for adjacency operators.

**Theorem 4.7** (Algebraic characterisation of strong connectivity for graphons). *Let  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  be an adjacency operator and  $\mathbb{L} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  be the corresponding graph-Laplacian operator defined as in (4.14). Then, if the underlying interaction topology is a disjoint union of strongly connected components in the sense of Definition 4.5, there exists an element  $v \in L^\infty(I, \mathbb{R}_+^*)$  such that*

$$\mathbb{L}^* v = 0, \quad \int_I v(i) di = 1 \quad \text{and} \quad \|v\|_{L^\infty(I, \mathbb{R}_+^*)} \leq \frac{1}{\delta},$$

where we write  $\mathbb{L}^* v = 0$  to implicitly mean that  $\mathbb{L}^*(v, \dots, v) = 0$  in  $L^2(I, \mathbb{R}^d)$ . If in addition one imposes that  $\int_{I_n} v(i) di = \mathcal{L}_{\sqcup I}^1(I_n)$  for every  $n \geq 1$ , then  $v \in L^\infty(I, \mathbb{R}_+^*)$  is uniquely determined.

*Proof.* In the case where  $d = 1$ , the results of [16, Theorem 1] yield the existence of a countable family of uniquely determined functions  $(v_n)_{n \geq 1} \subset L^2(I, \mathbb{R}_+^*)$  which satisfy

$$\text{supp}(v_n) = I_n, \quad \mathbb{L}_n^* v_n = 0 \quad \text{and} \quad \int_{I_n} v_n(i) di = \mathcal{L}_{\sqcup I}^1(I_n),$$

for each  $n \geq 1$ . Here,  $\mathbb{L}_n^* \in \mathcal{L}(L^2(I_n, \mathbb{R}))$  denotes the restriction of  $\mathbb{L}^*$  to  $L^2(I_n, \mathbb{R})$ , namely

$$(\mathbb{L}_n^* x)(i_n) = \int_{I_n} (a(i_n, j)x(i_n) - a(j, i_n)x(j)) dj \quad \text{for } \mathcal{L}_{\sqcup I}^1\text{-almost every } i_n \in I_n,$$

for each  $x \in L^2(I_n, \mathbb{R})$ . Then, one can show that the map defined by

$$v(i) := \sum_{n=1}^{+\infty} \mathbb{1}_{I_n}(i) v_n(i), \tag{4.17}$$

for  $\mathcal{L}_{\sqcup I}^1$ -almost every  $i \in I$  satisfies the conclusions of Theorem 4.7. Observe then that by (4.16), the fact that  $\mathbb{L}_n^* v_n = 0$  directly yields

$$v_n(i_n) = \frac{\int_{I_n} a(j, i) v_n(j) dj}{\int_{I_n} a(i_n, j) dj} \leq \frac{\int_{I_n} v_n(j) dj}{\delta \mathcal{L}_{\sqcup I}^1(I_n)} = \frac{1}{\delta},$$

for  $\mathcal{L}_{\sqcup I}^1$ -almost every  $i_n \in I_n$  and each  $n \geq 1$ . Because the sets  $(I_n)_{n \geq 1}$  form a disjoint partition of  $I$ , it follows easily from the definition of  $v$  in (4.17) that

$$\|v\|_{L^\infty(I, \mathbb{R}_+^*)} = \sup_{n \geq 1} \|v_n\|_{L^\infty(I_n, \mathbb{R}_+^*)} \leq \frac{1}{\delta}.$$

Finally when  $d \geq 1$  is arbitrary, the claims of Theorem 4.7 simply follow by defining the mapping  $\mathcal{V} : i \in I \mapsto (v(i), \dots, v(i)) \in \mathbb{R}^d$  with  $v \in L^\infty(I, \mathbb{R}_+^*)$  being given as in (4.17).  $\square$

Given an adjacency operator  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  defining a disjoint union of strongly connected components in the sense of Definition 4.5, we introduce the rescaled Laplacian operator

$$\mathbb{L}_v := \mathcal{M}_v \mathbb{L},$$

where the function  $v \in L^\infty(I, \mathbb{R}_+^*)$  appearing in the multiplication operator is the one given by Theorem 4.7. Then, we define the algebraic connectivity of the underlying interaction topology by

$$\lambda_2(\mathbb{L}_v) := \inf_{x \in \mathcal{E}^\perp} \frac{\langle \mathbb{L}_v x, x \rangle_{L^2(I, \mathbb{R}^d)}}{\|x\|_{L^2(I, \mathbb{R}^d)}^2} \geq 0, \tag{4.18}$$

and we will see in Section 4.2 that this quantity does allow to quantify consensus formation for strongly connected topologies. Therein, we will also prove that the rescaled Laplacian operator is indeed a positive semi-definite operator over  $L^2(I, \mathbb{R}^d)$ .

**Open problem.** *Is the converse implication of Theorem 4.1 still valid in the context of graphon models described by Theorem 4.7? Namely, is it true that an interaction topology is a countable disjoint union of strongly connected components whenever there exists an element  $v \in L^\infty(I, \mathbb{R}_+^*)$  such that  $\mathbb{L}^* v = 0$ ?*

## 4.2 Exponential consensus formation for algebraically persistent topologies

In this section, we investigate the asymptotic formation of  $L^2$ -consensus for graphon models whose interaction topologies are persistent in a suitable sense, quantified by means of the algebraic connectivity. We start by exposing in Section 4.2.1 a general convergence result for symmetric topologies whose time-averages have lower-bounded algebraic connectivity. We then proceed in Section 4.2.2 by studying consensus formation in balanced topologies, under the condition that the algebraic connectivity itself is persistent. We finally adapt the underlying strategy to study piecewise-constant (in time) disjoint unions of strongly connected components in Section 4.2.3.

### 4.2.1 Symmetric nonlinear interaction topologies

In what follows, we shall prove the following result, which describes the exponential formation of  $L^2$ -consensus for symmetric graphon dynamics, under the condition that the time-averages of the corresponding graph-Laplacian operators over time-windows of fixed length have a uniformly lower-bounded algebraic connectivity.

**Theorem 4.8** (Exponential  $L^2$ -consensus formation for symmetric topologies). *Fix an initial datum  $x^0 \in L^\infty(I, \mathbb{R}^d)$  with  $\|x^0\|_{L^\infty(I, \mathbb{R}^d)} \leq R$  for some  $R > 0$ , assume that hypotheses **(GD)** hold, and that  $a(t) \in L^\infty(I \times I, [0, 1])$  defines a symmetric interaction topology for  $\mathcal{L}^1$ -almost every  $t \geq 0$ . Moreover, suppose that there exists a pair of coefficients  $(\tau, \mu) \in \mathbb{R}_+^* \times (0, 1]$  such that the persistence condition*

$$\lambda_2 \left( \frac{1}{\tau} \int_t^{t+\tau} \mathbb{L}^a(s) ds \right) \geq \mu, \quad (4.19)$$

holds for all times  $t \geq 0$ , where  $\mathbb{L}^a(s) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is the linear part of the graph-Laplacian operator as defined in (4.10) for  $\mathcal{L}^1$ -almost every  $s \geq 0$ .

Then, there exist constants  $\alpha, \gamma > 0$  depending only on  $(\phi(\cdot), R, \tau, \mu)$  such that

$$\|x(t) - \bar{x}^0\|_{L^2(I, \mathbb{R}^d)} \leq \alpha \|x(0) - \bar{x}^0\|_{L^2(I, \mathbb{R}^d)} \exp(-\gamma \mu t), \quad (4.20)$$

for all times  $t \geq 0$ . In particular, the solutions of (4.9) exponentially converge to consensus in the  $L^2(I, \mathbb{R}^d)$ -norm topology.

Taking inspiration from some of the seminal works [19, 20] on the mathematical formalisation of consensus and flocking analysis for discrete symmetric multi-agent systems, we introduce the *variance bilinear form* over  $L^2(I, \mathbb{R}^d)$ , which is defined by

$$\mathbb{B}(x, y) := \int_I \langle x(i), y(i) \rangle di - \langle \bar{x}, \bar{y} \rangle, \quad (4.21)$$

for any  $x, y \in L^2(I, \mathbb{R}^d)$ , and we observe that the evaluation  $\mathbb{B}(x, x)$  coincides exactly with the  $L^2$ -distance between  $x \in L^2(I, \mathbb{R}^d)$  and the consensus manifold  $\mathcal{C}$ . Based on this observation, we consider for every solution  $x(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, L^2(I, \mathbb{R}^d))$  of (4.9) the *standard deviation map*

$$X(t) := \sqrt{\mathbb{B}(x(t), x(t))} = \|x(t) - \bar{x}(t)\|_{L^2(I, \mathbb{R}^d)}, \quad (4.22)$$

that is defined for all times  $t \geq 0$ . In the following proposition, we state some of the useful properties of the variance bilinear form, notably in conjunction with graph-Laplacians.

**Proposition 4.9** (Elementary properties of  $\mathbb{B}$ ). *The map  $x \in L^2(I, \mathbb{R}^d) \mapsto \mathbb{B}(x, x) \in \mathbb{R}_+$  defines a seminorm over  $L^2(I, \mathbb{R}^d)$ , which satisfies the Cauchy-Schwarz type inequality*

$$\mathbb{B}(x, y) \leq \sqrt{\mathbb{B}(x, x)} \sqrt{\mathbb{B}(y, y)}, \quad (4.23)$$

for each  $x, y \in L^2(I, \mathbb{R}^d)$ . In addition, let  $\mathbb{L} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  be the graph-Laplacian associated with an adjacency operator  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  defining a balanced topology in the sense of Definition 4.4. Then, one has that  $\mathbb{L}x \in \mathcal{C}^\perp$  as well as  $\mathbb{L}^*x \in \mathcal{C}^\perp$ , and the algebraic connectivity can be rewritten as

$$\lambda_2(\mathbb{L}) = \inf_{x \notin \mathcal{C}} \frac{\mathbb{B}(\mathbb{L}x, x)}{\mathbb{B}(x, x)} \geq 0. \quad (4.24)$$

*Proof.* Proving that  $\mathbb{B}(\cdot, \cdot)$  induces a seminorm over  $L^2(I, \mathbb{R}^d)$  is a matter of elementary computations. Concerning the Cauchy-Schwarz inequality, let it first be noted that

$$\mathbb{B}(x, y) = \mathbb{B}(x - \bar{x}, y) = \mathbb{B}(x, y - \bar{y}) = \mathbb{B}(x - \bar{x}, y - \bar{y}) = \langle x - \bar{x}, y - \bar{y} \rangle_{L^2(I, \mathbb{R}^d)},$$

for any  $x, y \in L^2(I, \mathbb{R}^d)$ . The inequality (4.23) then follows by an application of the standard Cauchy-Schwarz inequality for the scalar product of  $L^2(I, \mathbb{R}^d)$ . Let now  $\mathbb{L} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  be the graph-Laplacian operator associated with a balanced graphon, and notice that

$$\begin{aligned} \int_I (\mathbb{L}x)(i) di &= \int_I \int_I a(i, j)(x(i) - x(j)) dj di \\ &= \int_I \left( \int_I (a(i, j) - a(j, i)) dj \right) x(i) di = 0, \end{aligned} \quad (4.25)$$

for every  $x \in L^2(I, \mathbb{R}^d)$ , where we used Fubini's theorem along with Definition 4.4. This together with (4.12) implies that  $\mathbb{L}x \in \mathcal{C}^\perp$ , and one can show that  $\mathbb{L}^*x \in \mathcal{C}^\perp$  in the very same way. The alternative characterisation (4.24) of the algebraic connectivity simply follows from the observation that

$$\mathbb{B}(\mathbb{L}x, x) = \langle \mathbb{L}x, x \rangle_{L^2(I, \mathbb{R}^d)},$$

for any  $x \in L^2(I, \mathbb{R}^d)$ , as a consequence of (4.25). Finally, one can obtain the positive semi-definiteness of  $\mathbb{L}$  with respect to  $\mathbb{B}(\cdot, \cdot)$  by checking that

$$\begin{aligned} \mathbb{B}(\mathbb{L}x, x) &= \int_I \int_I a(i, j) \langle x(i), x(i) - x(j) \rangle dj di \\ &= \frac{1}{2} \int_I \int_I a(i, j) \left( |x(i) - x(j)|^2 + |x(i)|^2 - |x(j)|^2 \right) dj di \\ &= \frac{1}{2} \int_I \int_I a(i, j) |x(i) - x(j)|^2 dj di + \frac{1}{2} \int_I \left( \int_I a(i, j) dj - \int_I a(j, i) dj \right) |x(i)|^2 di \\ &= \frac{1}{2} \int_I \int_I a(i, j) |x(i) - x(j)|^2 dj di \geq 0, \end{aligned}$$

where we again used Fubini's theorem and the fact that the interaction topology is balanced.  $\square$

The proof of Theorem 4.8 will be conveyed by building a *strictly dissipative* Lyapunov function for (4.9), following a methodology developed in [14] for finite-dimensional symmetric multi-agent systems. Taking inspiration from the literature of strict Lyapunov design for persistent systems, we define given a trajectory  $x(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, L^2(I, \mathbb{R}^d))$  of (4.9) the time-dependent linear operator

$$\Psi_\tau(t) := (1 + c_\phi)\tau \text{Id} - \frac{1}{\tau} \int_t^{t+\tau} \int_t^s \mathbb{L}(\sigma, x(\sigma)) d\sigma ds \in \mathcal{L}(L^2(I, \mathbb{R}^d)), \quad (4.26)$$

for all times  $t \geq 0$ , where we recall that  $c_\phi = \sup_{r \in \mathbb{R}_+} \phi(r) < +\infty$  under hypotheses **(GD)**.

**Lemma 4.10** (Properties of the operators  $\Psi_\tau$ ). *The operators  $\Psi_\tau(t) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  are symmetric for all times  $t \geq 0$ , and satisfy the estimates*

$$\tau \mathbb{B}(y, y) \leq \mathbb{B}(\Psi_\tau(t)y, y) \leq (1 + c_\phi)\tau \mathbb{B}(y, y), \quad (4.27)$$

for each  $y \in L^2(I, \mathbb{R}^d)$ . Moreover, the map  $t \in \mathbb{R}_+ \mapsto \Psi_\tau(t) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is Lipschitz continuous and differentiable  $\mathcal{L}^1$ -almost everywhere, with

$$\frac{d}{dt} \Psi_\tau(t) = \mathbb{L}(t, x(t)) - \frac{1}{\tau} \int_t^{t+\tau} \mathbb{L}(s, x(s)) ds, \quad (4.28)$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ .

*Proof.* The fact that the operators  $\Psi_\tau(t) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  are symmetric for every  $t \geq 0$  is a simple consequence of the linearity of the Bochner integral. Given  $y \in L^2(I, \mathbb{R}^d)$ , one can show that

$$0 \leq \mathbb{B}(\mathbb{L}(\sigma, x(\sigma))y, y) \leq c_\phi \mathbb{B}(y, y), \quad (4.29)$$

where the first inequality follows from the positive semi-definiteness of  $\mathbb{L}(\sigma, x(\sigma))$  with respect to  $\mathbb{B}(\cdot, \cdot)$ , provided by Proposition 4.9 for  $\mathcal{L}^1$ -almost every  $\sigma \in [t, t + \tau]$ , while the second one can be obtained by remarking that

$$\begin{aligned} \mathbb{B}(\mathbb{L}(\sigma, x(\sigma))y, y) &= \frac{1}{2} \int_I \int_I a(\sigma, i, j) \phi(|x(\sigma, i) - x(\sigma, j)|) |y(i) - y(j)|^2 dj di \\ &\leq \frac{1}{2} \int_I \int_I c_\phi |y(i) - y(j)|^2 dj di \\ &= c_\phi \mathbb{B}(y, y), \end{aligned}$$

by symmetry of the weight functions  $a(\sigma) \in L^\infty(I \times I, [0, 1])$ . Upon integrating (4.29) with respect to  $\sigma \in [t, s]$  and then with respect to  $s \in [t, t + \tau]$  while merging the resulting expression with the definition (4.26) of  $\Psi_\tau(t) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$ , we obtain the lower- and upper-bounds

$$\tau \mathbb{B}(y, y) \leq \mathbb{B}(\Psi_\tau(t)y, y) \leq (1 + c_\phi) \tau \mathbb{B}(y, y),$$

for every  $y \in L^2(I, \mathbb{R}^d)$  and all times  $t \geq 0$ .

We now shift our attention to the regularity properties of the map  $t \in \mathbb{R}_+ \mapsto \Psi_\tau(t) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$ . Let  $t_2 \geq t_1 \geq 0$  be such that  $|t_2 - t_1| \leq \tau$ , and fix an element  $y \in L^2(I, \mathbb{R}^d)$ . Then, one has that

$$\begin{aligned} &\|(\Psi_\tau(t_2) - \Psi_\tau(t_1))y\|_{L^2(I, \mathbb{R}^d)} \\ &\leq \frac{1}{\tau} \int_{t_1}^{t_2} \int_{t_1}^s \|\mathbb{L}(\sigma, x(\sigma))y\|_{L^2(I, \mathbb{R}^d)} d\sigma ds + \frac{1}{\tau} \int_{t_2}^{t_1+\tau} \int_{t_1}^{t_2} \|\mathbb{L}(\sigma, x(\sigma))y\|_{L^2(I, \mathbb{R}^d)} d\sigma ds \\ &\quad + \frac{1}{\tau} \int_{t_1+\tau}^{t_2+\tau} \int_{t_2}^s \|\mathbb{L}(\sigma, x(\sigma))y\|_{L^2(I, \mathbb{R}^d)} d\sigma ds \\ &\leq (t_2 - t_1) \left( \frac{1}{\tau} \int_{t_1}^{t_2} \|\mathbb{L}(s, x(s))y\|_{L^2(I, \mathbb{R}^d)} ds + \frac{1}{\tau} \int_{t_2}^{t_1+\tau} \|\mathbb{L}(s, x(s))y\|_{L^2(I, \mathbb{R}^d)} ds \right. \\ &\quad \left. + \frac{1}{\tau} \int_{t_2}^{t_2+\tau} \|\mathbb{L}(s, x(s))y\|_{L^2(I, \mathbb{R}^d)} ds \right) \\ &\leq 2c_\phi |t_2 - t_1| \|y\|_{L^2(I, \mathbb{R}^d)}, \end{aligned}$$

where we used Fubini's theorem and the fact that  $\|\mathbb{L}(\sigma, x(\sigma))\|_{\mathcal{L}(L^2(I, \mathbb{R}^d))} \leq c_\phi$ . Therefore, we obtain

$$\|\Psi_\tau(t_2) - \Psi_\tau(t_1)\|_{\mathcal{L}(L^2(I, \mathbb{R}^d))} \leq 2c_\phi |t_2 - t_1|,$$

for each time instants  $t_2 \geq t_1 \geq 0$  satisfying  $|t_2 - t_1| \leq \tau$ . As the Lipschitz constant does not depend on  $t_1, t_2 \geq 0$ , we obtain that the map  $\Psi_\tau : \mathbb{R}_+ \rightarrow \mathcal{L}(L^2(I, \mathbb{R}^d))$  is Lipschitz continuous by covering the real line with intervals of length  $\tau$ . In particular, it is differentiable  $\mathcal{L}^1$ -almost everywhere as a consequence of the generalisation of Rademacher's theorem for maps with values in Hilbert spaces (see e.g. [42, Theorem 1.3]). Finally, the expression (4.28) of its time-derivative can be obtained by standard computations.  $\square$

Now that we have introduced these technical tools, we can move to the proof of Theorem 4.8.

*Proof of Theorem 4.8.* Let  $x \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, L^2(I, \mathbb{R}^d))$  be a solution of (4.9), and assume without loss of generality that  $x(t) \notin \mathcal{C}$  for all times  $t \geq 0$ . Given an arbitrary parameter  $\lambda > 0$  whose value will be fixed later on, we consider the candidate Lyapunov functional defined by

$$\mathcal{X}_\tau(t) := \lambda X(t) + \sqrt{\mathbb{B}(\Psi_\tau(t)x(t), x(t))}$$

for every  $t \geq 0$ , and it can be checked directly as a consequence of (4.27) that

$$(\lambda + \sqrt{\tau})X(t) \leq \mathcal{X}_\tau(t) \leq (\lambda + \sqrt{(1 + c_\phi)\tau})X(t) \tag{4.30}$$

for all times  $t \geq 0$ . Moreover, owing to the regularity of  $\Psi_\tau(\cdot)$ , the mapping  $t \in \mathbb{R}_+ \mapsto \mathcal{X}_\tau(t) \in \mathbb{R}_+$  is differentiable  $\mathcal{L}^1$ -almost everywhere, with

$$\begin{aligned} \frac{d}{dt} \mathcal{X}_\tau(t) &= -\frac{\lambda}{X(t)} \mathbb{B}(\mathbb{L}(t, x(t))x(t), x(t)) + \frac{\mathbb{B}\left(\frac{d}{dt} \Psi_\tau(t)x(t), x(t)\right) - 2\mathbb{B}(\Psi_\tau(t)x(t), \mathbb{L}(t, x(t))x(t))}{2\sqrt{\mathbb{B}(\Psi_\tau(t)x(t), x(t))}} \\ &= -\frac{\lambda}{X(t)} \mathbb{B}(\mathbb{L}(t, x(t))x(t), x(t)) - \frac{1}{2\sqrt{\mathbb{B}(\Psi_\tau(t)x(t), x(t))}} \mathbb{B}\left(\left(\frac{1}{\tau} \int_t^{t+\tau} \mathbb{L}(s, x(s)) ds\right)x(t), x(t)\right) \\ &\quad + \frac{1}{2\sqrt{\mathbb{B}(\Psi_\tau(t)x(t), x(t))}} \left( \mathbb{B}(\mathbb{L}(t, x(t))x(t), x(t)) - 2\mathbb{B}(\Psi_\tau(t)x(t), \mathbb{L}(t, x(t))x(t)) \right), \end{aligned}$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ , where we used the fact that  $\Psi_\tau(t)$  is symmetric. In addition, as a consequence of the persistence condition (4.19), one has that

$$\begin{aligned} &\mathbb{B}\left(\left(\frac{1}{\tau} \int_t^{t+\tau} \mathbb{L}(s, x(s)) ds\right)x(t), x(t)\right) \\ &= \frac{1}{2} \int_I \int_I \left(\frac{1}{\tau} \int_t^{t+\tau} a(s, i, j) \phi(|x(s, i) - x(s, j)|) ds\right) |x(t, i) - x(t, j)|^2 dj di \\ &\geq \frac{1}{2} \int_I \int_I \gamma_R \left(\frac{1}{\tau} \int_t^{t+\tau} a(s, i, j) ds\right) |x(t, i) - x(t, j)|^2 dj di \\ &= \gamma_R \mathbb{B}\left(\left(\frac{1}{\tau} \int_t^{t+\tau} \mathbb{L}^a(s) ds\right)x(t), x(t)\right) \\ &\geq \gamma_R \mu X(t)^2, \end{aligned}$$

where  $\gamma_R := \min_{r \in [0, 2R]} \phi(r)$ . This latter expression along with the positive semi-definiteness of  $\mathbb{L}(t, x(t))$  with respect to  $\mathbb{B}(\cdot, \cdot)$  and the bounds of (4.27) further yields

$$\begin{aligned} \frac{d}{dt} \mathcal{X}_\tau(t) &\leq -\frac{\gamma_R \mu}{2\sqrt{(1+c_\phi)\tau}} X(t) + \frac{1}{X(t)} \left( \frac{1}{\sqrt{\tau}} - \sqrt{(1+c_\phi)\tau} - \lambda \right) \mathbb{B}(\mathbb{L}(t, x(t))x(t), x(t)) \\ &\quad + \frac{1}{\sqrt{\mathbb{B}(\Psi_\tau(t)x(t), x(t))}} \mathbb{B}\left(\left(\frac{1}{\tau} \int_t^{t+\tau} \int_t^s \mathbb{L}(\sigma, x(\sigma)) d\sigma ds\right)x(t), \mathbb{L}(t, x(t))x(t)\right), \end{aligned} \tag{4.31}$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ . We now need to estimate the last term in the right-hand side of the previous inequality. As consequence of the Cauchy-Schwarz inequality supported by  $\mathbb{B}(\cdot, \cdot)$  and Young's inequality, it holds for every  $\varepsilon > 0$  that

$$\begin{aligned} &\mathbb{B}\left(\left(\frac{1}{\tau} \int_t^{t+\tau} \int_t^s \mathbb{L}(\sigma, x(\sigma)) d\sigma ds\right)x(t), \mathbb{L}(t, x(t))x(t)\right) \\ &\leq \sqrt{\mathbb{B}\left(\frac{1}{\tau} \int_t^{t+\tau} \int_t^s \mathbb{L}(\sigma, x(\sigma)) d\sigma ds\right)x(t), \frac{1}{\tau} \int_t^{t+\tau} \int_t^s \mathbb{L}(\sigma, x(\sigma)) d\sigma ds\right) \sqrt{\mathbb{B}(\mathbb{L}(t, x(t))x(t), \mathbb{L}(t, x(t))x(t))} \\ &\leq \left\| \frac{1}{\tau} \int_t^{t+\tau} \int_t^s \mathbb{L}(\sigma, x(\sigma)) d\sigma ds \right\|_{\mathbb{B}}^{1/2} X(t) \left( \sqrt{\mathbb{B}(\mathbb{L}(t, x(t))x(t), \mathbb{L}(t, x(t))x(t))} \right) \\ &\leq \frac{\varepsilon}{2} \left\| \frac{1}{\tau} \int_t^{t+\tau} \int_t^s \mathbb{L}(\sigma, x(\sigma)) d\sigma ds \right\|_{\mathbb{B}} X(t)^2 + \frac{1}{2\varepsilon} \mathbb{B}(\mathbb{L}(t, x(t))x(t), \mathbb{L}(t, x(t))x(t)), \end{aligned} \tag{4.32}$$

where  $\|\cdot\|_{\mathbb{B}}$  denotes the operator semi-norm induced by  $\mathbb{B}(\cdot, \cdot)$  on  $\mathcal{L}(L^2(I, \mathbb{R}^d))$ . By Jensen's inequality, one can further estimate the latter as

$$\left\| \frac{1}{\tau} \int_t^{t+\tau} \int_t^s \mathbb{L}(\sigma, x(\sigma)) d\sigma ds \right\|_{\mathbb{B}} \leq \frac{1}{\tau} \int_t^{t+\tau} \int_t^s \|\mathbb{L}(\sigma, x(\sigma))\|_{\mathbb{B}} d\sigma ds \leq c_\phi \tau, \tag{4.33}$$

since  $\|\mathbb{L}(\sigma, x(\sigma))\|_{\mathbb{B}} \leq c_\phi$  for  $\mathcal{L}^1$ -almost every  $\sigma \in [t, t+\tau]$  owing to (4.29). By the square-root lemma for positive semi-definite symmetric operators (see e.g. [69, Theorem VI.9]), it also holds that

$$\begin{aligned} \mathbb{B}(\mathbb{L}(t, x(t))x(t), \mathbb{L}(t, x(t))x(t)) &\leq \sqrt{c_\phi} \mathbb{B}\left((\mathbb{L}(t, x(t))^* \mathbb{L}(t, x(t)))^{1/2} x(t), x(t)\right) \\ &= \sqrt{c_\phi} \mathbb{B}(\mathbb{L}(t, x(t))x(t), x(t)), \end{aligned} \tag{4.34}$$

by the symmetry of  $\mathbb{L}(t, x(t))$  for  $\mathcal{L}^1$ -almost every  $t \geq 0$ . Hence, by merging the estimates of (4.32), (4.33) and (4.34) into the differential inequality (4.31), one recovers that

$$\begin{aligned} \frac{d}{dt} \mathcal{X}_\tau(t) &\leq \left( \frac{\varepsilon}{2\sqrt{\tau}} - \frac{\gamma_R \mu}{2\sqrt{(1+c_\phi)\tau}} \right) X(t) \\ &\quad + \frac{1}{X(t)} \left( \frac{1}{\sqrt{\tau}} + \frac{1}{2\sqrt{\tau\varepsilon}} - \sqrt{(1+c_\phi)\tau} - \lambda \right) \mathbb{B}(\mathbb{L}(t, x(t))x(t), x(t)), \end{aligned} \quad (4.35)$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ .

Our goal now is to choose the free parameters  $\lambda, \varepsilon > 0$  in such a way that (4.35) yields a strictly dissipative estimate on the standard deviation  $X(\cdot)$ . To this end, we fix

$$\varepsilon := \frac{\gamma_R \mu}{2\sqrt{(1+c_\phi)}} \quad \text{and} \quad \lambda := \frac{1}{\sqrt{\tau}} + \frac{1}{2\sqrt{\tau\varepsilon}} - \sqrt{(1+c_\phi)\tau}, \quad (4.36)$$

which together with the estimates of (4.30) allows us to rewrite (4.35) as

$$\frac{d}{dt} \mathcal{X}_\tau(t) \leq - \frac{\mu}{4\sqrt{(1+c_\phi)\tau}(\lambda + \sqrt{(1+c_\phi)\tau})} \mathcal{X}_\tau(t). \quad (4.37)$$

By an application of Grönwall's lemma together with yet another estimation using the bounds of (4.30), we can conclude that

$$X(t) \leq \left( \lambda + \sqrt{(1+c_\phi)\tau} \right) X(0) \exp \left( - \frac{\mu}{4\sqrt{(1+c_\phi)\tau}(\lambda + \sqrt{(1+c_\phi)\tau})} t \right),$$

for all times  $t \geq 0$ . Recalling the definition (4.22) of the standard deviation and defining the constants

$$\alpha := \left( \lambda + \sqrt{(1+c_\phi)\tau} \right) \quad \text{and} \quad \gamma := \frac{1}{4\sqrt{(1+c_\phi)\tau}(\lambda + \sqrt{(1+c_\phi)\tau})},$$

with  $\lambda > 0$  being given as in (4.36), we obtain the exponential estimate claimed in (4.20), which concludes the proof of Theorem 4.8.  $\square$

## 4.2.2 Balanced linear interactions topologies

In this second section devoted to the investigation of  $L^2$ -consensus formation, we prove the following theorem, which shows that linear balanced graphon dynamics exponentially converge to consensus when the underlying algebraic connectivity is persistent. Throughout this section, we assume that there are no nonlinear interactions between agents – namely  $\phi(\cdot) \equiv 1$  –, so that  $\mathbb{L}(t, x) = \mathbb{L}(t)$  for  $\mathcal{L}^1$ -almost every  $t \geq 0$  and any  $x \in L^2(I, \mathbb{R}^d)$ .

**Theorem 4.11** (Exponential  $L^2$ -consensus formation for balanced topologies). *Fix an initial datum  $x^0 \in L^2(I, \mathbb{R}^d)$ , assume that hypothesis (GD)-(i) holds, and that the interaction kernel  $t \in \mathbb{R}_+ \mapsto a(t) \in L^\infty(I \times I, [0, 1])$  defines balanced interaction topologies in the sense of Definition 4.4 for  $\mathcal{L}^1$ -almost every  $t \geq 0$ . Moreover, suppose that there exists a pair of coefficients  $(\tau, \mu) \in \mathbb{R}_+^* \times (0, 1]$  such that the persistence condition*

$$\frac{1}{\tau} \int_t^{t+\tau} \lambda_2(\mathbb{L}(s)) ds \geq \mu, \quad (4.38)$$

holds for all times  $t \geq 0$ .

Then, there exists a constant  $\alpha > 0$  depending only on  $(\tau, \mu)$  such that

$$\|x(t) - \bar{x}^0\|_{L^2(I, \mathbb{R}^d)} \leq \alpha \|x(0) - \bar{x}^0\|_{L^2(I, \mathbb{R}^d)} \exp(-\mu t), \quad (4.39)$$

for all times  $t \geq 0$ . In particular, the solutions of (4.9) exponentially converge to consensus in the  $L^2(I, \mathbb{R}^d)$ -norm topology.

*Proof.* Similarly to what was done in Section 4.2.1 for symmetric graphon models, we will derive dissipative decay estimates on the standard deviation map  $X(\cdot)$  defined along solutions of (4.9) as in (4.22). For balanced dynamics, it can be easily checked that

$$\frac{d}{dt}\bar{x}(t) = - \int_I (\mathbb{L}(t)x(t))(i)di = 0,$$

as a consequence of Proposition 4.9, which implies that  $\bar{x}(t) = \bar{x}^0$  for all times  $t \geq 0$ . Thus, the time-derivative of the standard deviation can be written as

$$\frac{d}{dt}X(t) = -\frac{1}{X(t)}\mathbb{B}(\mathbb{L}(t)x(t), x(t)) \leq -\lambda_2(\mathbb{L}(t))X(t), \quad (4.40)$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ , where we assumed without loss of generality that  $x(t) \notin \mathcal{C}$ . By applying Grönwall's lemma and adapting the computations of (3.17) in the proof of Theorem 3.6, we obtain

$$X(t) \leq X(0) \exp(-\mu(t - \tau)),$$

for all times  $t \geq 0$ . Observing finally that  $X(t) = \|x(t) - \bar{x}^0\|_{L^2(I, \mathbb{R}^d)}$  in our context, we recover the exponential convergence to consensus (4.39) with  $\alpha := \exp(\mu\tau)$ .  $\square$

It is important to notice that the persistence conditions imposed in Theorem 4.8 and Theorem 4.11 are very different in nature. While the latter, written in (4.38), expresses the fact that  $\lambda_2(\mathbb{L}(\cdot))$  is positive sufficiently often on each time window of length  $\tau > 0$ , the former merely requires that the algebraic connectivity of the *average* topology is lower-bounded. In particular, it is perfectly possible to build time-dependent signals satisfying (4.19) for which  $\lambda_2(\mathbb{L}(\cdot)) \equiv 0$  (see e.g. [14, Section 4] for a finite-dimensional example, as well as Section 5 below). This practical and heuristic comparison between (4.19) and (4.38) can be made rigorous by means of the following general result.

**Proposition 4.12** (Concavity of the algebraic connectivity). *The algebraic connectivity defined as in (4.15) and seen as a mapping  $\lambda_2 : \mathcal{L}(L^2(I, \mathbb{R}^d)) \rightarrow \mathbb{R}$  is concave. In particular, given a time-dependent operator  $T \in L^2_{\text{loc}}(\mathbb{R}_+, \mathcal{L}(L^2(I, \mathbb{R}^d)))$  and a coefficient  $\tau \in \mathbb{R}_+^*$ , it holds that*

$$\frac{1}{\tau} \int_t^{t+\tau} \lambda_2(T(s))ds \leq \lambda_2\left(\frac{1}{\tau} \int_t^{t+\tau} T(s)ds\right), \quad (4.41)$$

for all times  $t \geq 0$ .

*Proof.* Let  $\zeta \in [0, 1]$  be a fixed parameter,  $T_1, T_2 \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  be two bounded operators and  $\varepsilon > 0$  be given. By the definition (4.15) of the algebraic connectivity, there exists  $x_\varepsilon \in \mathcal{C}^\perp$  such that

$$\frac{\langle ((1 - \zeta)T_1 + \zeta T_2)x_\varepsilon, x_\varepsilon \rangle_{L^2(I, \mathbb{R}^d)}}{\|x_\varepsilon\|_{L^2(I, \mathbb{R}^d)}^2} \leq \lambda_2\left((1 - \zeta)T_1 + \zeta T_2\right) + \varepsilon. \quad (4.42)$$

Because  $x_\varepsilon \in \mathcal{C}^\perp$ , it also holds

$$\lambda_2(T_1) \leq \frac{\langle T_1 x_\varepsilon, x_\varepsilon \rangle_{L^2(I, \mathbb{R}^d)}}{\|x_\varepsilon\|_{L^2(I, \mathbb{R}^d)}^2} \quad \text{and} \quad \lambda_2(T_2) \leq \frac{\langle T_2 x_\varepsilon, x_\varepsilon \rangle_{L^2(I, \mathbb{R}^d)}}{\|x_\varepsilon\|_{L^2(I, \mathbb{R}^d)}^2}, \quad (4.43)$$

and by combining the estimates of (4.42) and (4.43), we recover the inequality

$$(1 - \zeta)\lambda_2(T_1) + \zeta\lambda_2(T_2) \leq \lambda_2\left((1 - \zeta)T_1 + \zeta T_2\right) + \varepsilon,$$

which holds for every  $\zeta \in [0, 1]$  and each  $T_1, T_2 \in \mathcal{L}(L^2(I, \mathbb{R}^d))$ . Since  $\varepsilon > 0$  is arbitrary, we conclude that  $\lambda_2 : \mathcal{L}(L^2(I, \mathbb{R}^d)) \rightarrow \mathbb{R}$  is concave. Given a curve  $T \in L^2(\mathbb{R}_+, \mathcal{L}(I, \mathbb{R}^d))$  of operators, the identity (4.41) simply follows from an application of the measure-theoretic version of Jensen's inequality (see e.g. [1, Lemma 1.15]).  $\square$

**Open problem** (Exponential consensus for algebraically persistent balanced topologies). *Do the conclusions of Theorem 4.11 still hold if one replaces the pointwise persistence condition (4.38) on the algebraic connectivity by the more general one (4.19) appearing in Theorem 4.8 ?*



### 4.2.3 Disjoint unions of strongly connected linear topologies

In this third and last section dealing with the formation of  $L^2$ -consensus, we prove that a result similar to Theorem 4.11 holds for piecewise constant switching topologies describing disjoint unions of strongly connected components. As in Section 4.2.2 above, we make the assumption that  $\phi(\cdot) \equiv 1$ , so that  $\mathbb{L}(t, x) = \mathbb{L}(t)$  for  $\mathcal{L}^1$ -almost every  $t \geq 0$  and any  $x \in L^2(I, \mathbb{R}^d)$ .

In what follows, we assume that there exists an increasing sequence  $(t_k)_{k \geq 0} \subset \mathbb{R}_+$ , satisfying

$$t_0 = 0 \quad \text{and} \quad t_{k+1} - t_k \geq \tau_d, \quad (4.44)$$

for a given *dwell-time*  $\tau_d > 0$ , and such that  $t \in [t_k, t_{k+1}) \mapsto a(t) \in L^\infty(I \times I, [0, 1])$  is constant for each  $k \geq 0$ . We point to the reference monograph [48, Section 3.2] for a precise discussion of the role of dwell-times in the stability of switched systems.

**Theorem 4.13** (Exponential  $L^2$ -consensus formation for strongly connected topologies). *Fix an initial datum  $x^0 \in L^2(I, \mathbb{R}^d)$ , assume that hypothesis (GD)-(i) holds and that the signal  $t \in \mathbb{R}_+ \mapsto a(t) \in L^\infty(I \times I, [0, 1])$  is piecewise constant and satisfies the dwell-time condition (4.44). Moreover, suppose that  $a(t)$  defines a disjoint union of strongly connected components in the sense of Definition 4.4 for  $\mathcal{L}^1$ -almost every  $t \geq 0$ . Denoting by  $v(t) \in L^\infty(I, \mathbb{R}_+^*)$  the canonical vector associated with  $a(t) \in L^\infty(I \times I, [0, 1])$  via Theorem 4.7, we further assume that there exists a triplet of coefficients  $(\tau, \mu, \nu) \in \mathbb{R}_+^* \times (0, 1]$  satisfying the inequality*

$$\frac{2}{\mu\nu^2} \log\left(\frac{1}{\nu}\right) < \tau_d, \quad (4.45)$$

and which are such that

$$\frac{1}{\tau} \int_t^{t+\tau} \lambda_2(\mathbb{L}_v(s)) ds \geq \mu, \quad (4.46)$$

and

$$\nu \leq v(t, i) \leq \frac{1}{\nu} \quad \text{for } \mathcal{L}^1_I\text{-almost every } i \in I, \quad (4.47)$$

for all times  $t \geq 0$ . Here  $\mathbb{L}_v(s) := \mathcal{M}_{v(s)}\mathbb{L}(s) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  and the algebraic connectivity  $\lambda_2(\mathbb{L}_v(s))$  is defined as in (4.18) for  $\mathcal{L}^1$ -almost every  $s \in [t, t + \tau]$ .

Then, there exists an element  $x^\infty \in \mathcal{C}(x^0)$  and a constant  $\alpha > 0$  depending only on  $(\tau, \mu, \nu)$  such that

$$\|x(t) - x^\infty\|_{L^2(I, \mathbb{R}^d)} \leq \alpha \|x(0) - \bar{x}^0\|_{L^2(I, \mathbb{R}^d)} \exp\left(-\left(\mu\nu^2 - \frac{2}{\tau_d} \log\left(\frac{1}{\nu}\right)\right)t\right), \quad (4.48)$$

for all times  $t \geq 0$ . In particular, the solutions of (4.9) exponentially converge to consensus in the  $L^2(I, \mathbb{R}^d)$ -norm topology.

A discussion shedding light on the lower-bound imposed on  $v(t)$  in (4.47) is proposed in Remark 5.4, based on the results of Section 5.1.

*Proof of Theorem 4.13.* Let it first be noted that  $\mathbb{L}_v(t)x \in \mathcal{C}^\perp$  for  $\mathcal{L}^1$ -almost every  $t \geq 0$  and each  $x \in L^2(I, \mathbb{R}^d)$ , since

$$\begin{aligned} \int_I (\mathbb{L}_v(t)x)(i) di &= \int_I \int_I v(t, i) a(t, i, j) (x(i) - x(j)) dj di \\ &= \int_I \left( \int_I (a(t, i, j) v(t, i) - a(t, j, i) v(t, j)) dj \right) x(i) di = 0, \end{aligned} \quad (4.49)$$

by Fubini's theorem along with the fact that  $\mathbb{L}(t)^*v(t) = 0$ . Therefore, the weighted barycenters

$$\bar{x}_v : t \in \mathbb{R}_+ \mapsto \int_I v(t, i) x(t, i) di \in \mathbb{R}^d,$$

associated with  $v(\cdot)$  are constant on each time interval of the form  $[t_k, t_{k+1})$  with  $k \geq 0$ . Following the proof strategy developed for Theorem 4.11, we introduce weighted variance bilinear forms

$$\mathbb{B}_{v(t)}(x, y) := \langle \mathcal{M}_{v(t)} x, y \rangle_{L^2(I, \mathbb{R}^d)} - \langle \bar{x}_{v(t)}, \bar{y}_{v(t)} \rangle,$$

defined for each  $x, y \in L^2(I, \mathbb{R}^d)$  and  $\mathcal{L}^1$ -almost every  $t \geq 0$ . In this context, one can check that

$$\mathbb{B}_{v(t)}(x(t) - \bar{x}_v(t), x(t) - \bar{x}_v(t)) = \mathbb{B}_{v(t)}(x(t), x(t)),$$

where both equalities rely on the fact that  $\int_I v(t, i) di = 1$  for all times  $t \geq 0$  by construction. Similarly, one can again use (4.49) to verify that

$$\langle \mathbb{L}_v(t)x, x \rangle_{L^2(I, \mathbb{R}^d)} = \mathbb{B}_{v(t)}(\mathbb{L}(t)x, x),$$

for each  $x \in L^2(I, \mathbb{R}^d)$  and  $\mathcal{L}^1$ -almost every  $t \geq 0$ , so that the algebraic connectivity defined in (4.18) can be rewritten as

$$\lambda_2(\mathbb{L}_v(t)) = \inf_{x \notin \mathcal{C}} \frac{\mathbb{B}_{v(t)}(\mathbb{L}(t)x, x)}{\mathbb{B}(x, x)} \geq 0.$$

The inequality on the right comes from the fact that  $\mathbb{L}(t)$  is positive semi-definite with respect to  $\mathbb{B}_{v(t)}(\cdot, \cdot)$ , which can be shown by a small modification of the argument in the proof of Proposition 4.9. Finally, we introduce the piecewise constant weighted standard deviations, which are defined by

$$X_v(t) := \sqrt{\mathbb{B}_{v(t)}(x(t), x(t))},$$

for all times  $t \geq 0$ . Using again the fact that  $\int_I v(t, i) di = 1$  for all times  $t \geq 0$ , one can show via routine computations that the weighted standard deviation can be alternatively expressed as

$$X_v(t) = \left( \int_I \int_I v(t, i) v(t, j) |x(t, i) - x(t, j)|^2 dj di \right)^{1/2}.$$

Together with the bounds (4.47) imposed on  $v(t) \in L^\infty(I, \mathbb{R}_+^*)$ , this allows us to recover the estimates

$$\nu X(t) \leq X_v(t) \leq \frac{1}{\nu} X(t), \quad (4.50)$$

for all times  $t \geq 0$ , where  $X(\cdot)$  stands for the usual standard deviation defined as in (4.22).

Observe now that for each  $k \geq 0$ , the map  $X_v(\cdot)$  is differentiable over  $[t_k, t_{k+1})$ , with

$$\frac{d}{dt} X_v(t) = -\frac{1}{X_v(t)} \mathbb{B}_{v(t)}(\mathbb{L}(t)x(t), x(t)) \leq -\frac{1}{X_v(t)} \lambda_2(\mathbb{L}_v(t)) X(t)^2 \leq -\nu^2 \lambda_2(\mathbb{L}_v(t)) X_v(t). \quad (4.51)$$

Then, by a simple application of Grönwall's lemma, it follows that

$$X_v(t) \leq X_v(t_k) \exp\left(-\nu^2 \lambda_2(\mathbb{L}_v(t_k))(t - t_k)\right), \quad (4.52)$$

for all times  $t \in [t_k, t_{k+1})$  and each  $k \geq 0$ . It is worth noting that  $X_v(\cdot)$  is in fact a discontinuous Lyapunov function, so that (4.52) cannot be used directly to prove the convergence of the system to consensus. By combining (4.52) and (4.50), we obtain the dissipation inequality

$$X(t) \leq \frac{1}{\nu} X(t_k) \exp\left(-\nu^2 \lambda_2(\mathbb{L}_v(t_k))(t - t_k)\right),$$

which holds for all times  $t \in [t_k, t_{k+1})$  and each  $k \geq 0$ . Thus, by performing a simple induction argument and repeating the computations in (3.17) involving the persistence condition (4.46), we can in turn recover that

$$X(t) \leq X(0) \frac{1}{\nu^{2(k+1)}} \exp\left(-\nu \int_0^t \lambda_2(\mathbb{L}_v(s)) ds\right) \leq \frac{\exp(\mu \nu^2 \tau)}{\nu^2} X(0) \frac{\exp(-\mu \nu^2 t)}{\nu^{2t/\tau_d}}, \quad (4.53)$$

for all  $t \in [t_k, t_{k+1})$ , where we used the fact that  $t_k \geq k\tau_d$  for each  $k \geq 0$  as a consequence of the dwell-time assumption (4.44). Finally, owing to the constraint (4.45) imposed on the dwell-time and connectivity coefficients, it holds in particular that

$$X(t) \leq \frac{\exp(\mu \nu^2 \tau)}{\nu^2} X(0) \exp\left(-\left(\mu \nu^2 - \frac{2}{\tau_d} \log\left(\frac{1}{\nu}\right)\right)t\right) \xrightarrow{t \rightarrow +\infty} 0, \quad (4.54)$$

so that the standard deviation  $X(\cdot)$  exponentially converges to zero as  $t \rightarrow +\infty$ .

Our goal now is to show that  $x(\cdot)$  converges to consensus. To this end, consider two arbitrary instants  $t_2 \geq t_1 \geq 0$ , and observe that the classical barycenters of the curve  $x(\cdot)$  are such that

$$\begin{aligned} |\bar{x}(t_2) - \bar{x}(t_1)| &\leq \int_I |x(t_2, i) - x(t_1, i)| di \\ &\leq \int_{t_1}^{t_2} \left( \int_I \int_I a(s, i, j) |x(s, j) - x(s, i)| dj di \right) ds \leq \int_{t_1}^{+\infty} X(s) ds \xrightarrow{t_1 \rightarrow +\infty} 0, \end{aligned} \quad (4.55)$$

where we used Fubini's theorem and Cauchy-Schwarz's inequality along with the decay estimate of (4.54). This implies that the curve  $\bar{x}(\cdot) \in C^0(\mathbb{R}_+, \mathbb{R}^d)$  converges exponentially to some limit point  $x^\infty \in \mathbb{R}^d$  as  $t \rightarrow +\infty$ . By combining (4.54) and (4.55), we can conclude that

$$\|x(t) - x^\infty\|_{L^2(I, \mathbb{R}^d)} \leq X(t) + \|\bar{x}(t) - x^\infty\|_{L^2(I, \mathbb{R}^d)} \leq \alpha X(0) \exp\left(-(\mu\nu^2 - \frac{2}{\tau_d} \log(\frac{1}{\nu}))t\right),$$

for a given constant  $\alpha > 0$  that only depends on the coefficients  $(\tau, \mu, \nu)$ , so that solutions of (4.9) exponentially converge to consensus in the  $L^2(I, \mathbb{R}^d)$ -norm topology.  $\square$

**Open problem** (Asymptotic consensus formation without dwell-times). *Would a weaker form of Theorem 4.13 still hold for general time-varying signals that do not satisfy a dwell-time condition?*

## 5 Further discussions around connectivity and consensus formation

In what follows, we elaborate on the properties of the algebraic connectivity defined in Section 4.1, as well as on its interplay with consensus formation. In Section 5.1, we prove that interaction topologies which are either symmetric or written as disjoint unions of strongly connected components are strongly connected if and only if their algebraic connectivity is positive. Then, in Section 5.2, we prove that the formation of  $L^2$ - and  $L^\infty$ -consensus are equivalent when the in-degree function is persistent in a suitable sense, and close the manuscript by a series of numerical illustrations presented in Section 5.3

### 5.1 Link between algebraic and graphon connectivity

In Section 4.2, we have shown that the notion of algebraic connectivity proposed in Section 4.1 allows for a systematic investigation of consensus formations for various kinds of time-dependent graphon models. As recalled above, the classical notion of algebraic connectivity encodes important structural properties of the underlying graph, including information about its connectivity. Our aim in this section is to discuss the counterpart of these results for graphon models.

We start by a technical lemma concerning the existence of a largest strongly connected component containing a point in a graphon. In the sequel, for the sake of conciseness, we shall say that two points  $i, j \in I$  are *m-connected* if  $i$  can be linked to  $j$  and reciprocally  $j$  can be linked to  $i$  in the sense of Definition 4.5-(a) by directed chains of Lebesgue points of length at most  $m \geq 1$ .

**Lemma 5.1** (Existence of largest strongly connected components). *Let  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  and  $i \in I$  be a Lebesgue point of the interaction kernel  $a \in L^\infty(I \times I, [0, 1])$ . Then, there exists a strongly connected set  $\mathcal{S}(i) \subset I$  in the sense of Definition 4.5-(a) that contains  $i$ , and which is maximal in the sense that*

$$\mathcal{L}_{\perp I}^1(J \setminus \mathcal{S}(i)) = 0,$$

for any other strongly connected set  $J \subset I$  containing  $\mathcal{S}(i)$ .

*Proof.* Given  $N \geq 1$ , let  $\mathcal{S}_N(i) \in \mathcal{P}(I)$  be defined as

$$\mathcal{S}_N(i) := \left\{ j \in I \text{ s.t. } i \text{ and } j \text{ are } m\text{-connected for some } m \leq N \right\},$$

and consider the union over  $N \geq 1$  of these sets

$$\mathcal{S}(i) := \bigcup_{N \geq 1} \mathcal{S}_N(i).$$

By construction, it is clear that  $\mathcal{S}(i)$  is a strongly connected subset of  $I$  in the sense of Definition 4.5-(a). Let  $J \subset I$  be a strongly connected set containing  $\mathcal{S}(i)$  and suppose that  $\mathcal{L}_{\perp I}^1(J \setminus \mathcal{S}(i)) > 0$ , which can be equivalently rewritten as

$$\mathcal{L}_{\perp I}^1\left(J \setminus \left(\bigcup_{N \geq 1} \mathcal{S}_N(i)\right)\right) = \mathcal{L}_{\perp I}^1\left(\bigcap_{N \geq 1} (J \setminus \mathcal{S}_N(i))\right) = \lim_{N \rightarrow +\infty} \mathcal{L}_{\perp I}^1(J \setminus \mathcal{S}_N(i)) > 0, \quad (5.1)$$

since the sequence of sets  $(\mathcal{S}_N(i))_{N \geq 1} \subset \mathcal{P}(I)$  is increasing. Then for every  $N \geq 1$  large enough, the identity (5.1) implies that there exists a subset  $J_N \subset J$  satisfying  $\mathcal{L}_{\perp I}^1(J_N) \geq \mathcal{L}_{\perp I}^1(J \setminus \mathcal{S}(i)) > 0$ , whose elements cannot be connected to  $i$  by a path of length at most  $N \geq 1$ . This contradicts our primary assumption that  $J$  is strongly connected, and shows that  $\mathcal{S}(i)$  is maximal.  $\square$

This technical result being established, we can provide a first correspondence between the positivity of the algebraic connectivity and graphon connectivity for symmetric interaction topologies.

**Theorem 5.2** (On the connectivity of undirected graphons). *Let  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  be a symmetric adjacency operator and  $\mathbb{L} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  be the corresponding graph-Laplacian. Then, the interaction topology described by  $\mathcal{A}$  is strongly connected in the sense of Definition 4.5 if and only if  $\lambda_2(\mathbb{L}) > 0$ .*

*Proof. Converse implication.* We start by assuming that  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is an adjacency operator for which  $\lambda_2(\mathbb{L}) > 0$ , and we will prove that condition (b) of Definition 4.5 holds, namely  $\inf_{i \in I} \int_I a(i, j) dj > 0$ . We recall that the in-degree function  $d \in L^\infty(I, [0, 1])$  associated with  $\mathcal{A}$  writes

$$d(i) := \int_I a(i, j) dj,$$

for  $\mathcal{L}_{\perp I}^1$ -almost every  $i \in I$ , and that  $\sigma_{\text{ess}}(\mathcal{M}_d) = \text{rg}(d)$  as a consequence of Proposition 2.6. Moreover, let it be noted that

$$\sigma_{\text{ess}}(\mathbb{L}) = \sigma_{\text{ess}}(\mathcal{M}_d - \mathcal{A}) = \sigma_{\text{ess}}(\mathcal{M}_d),$$

because  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is a compact operator (see e.g. [69, Section VI.5]), and the essential spectrum is invariant under compact perturbations (see e.g. [68, Example XIII.4.3]). Then, by the min-max theorem for self-adjoint operators (see e.g. [68, Theorem XIII.1]), we obtain the lower-bound

$$0 < \lambda_2(\mathbb{L}) \leq \inf \left\{ \lambda \in \mathbb{R} \text{ s.t. } \lambda \in \sigma_{\text{ess}}(\mathbb{L}) \right\},$$

on the essential spectrum of the graph-Laplacian operator  $\mathbb{L}$ . In particular, it necessarily holds that  $\inf_{i \in I} d(i) = \inf_{i \in I} \int_I a(i, j) dj > 0$  whenever  $\lambda_2(\mathbb{L}) > 0$ .

We turn our attention to condition (a) of Definition 4.5. Assume by contradiction that it is violated, namely that the disconnection set

$$\mathcal{D} := \left\{ (i, j) \in I \times I \text{ s.t. } i \text{ and } j \text{ are not } m\text{-connected for each } m \geq 1 \right\}, \quad (5.2)$$

has positive  $\mathcal{L}_{\perp I \times I}^2$ -measure. We claim that in this case, there exists a subset  $\mathcal{S} \subset I$  such that

$$\mathcal{L}_{\perp I}^1(\mathcal{S}) > 0, \quad \mathcal{L}_{\perp I}^1(I \setminus \mathcal{S}) > 0 \quad \text{and} \quad a(i, j) = 0 \quad \text{for } \mathcal{L}_{\perp I \times I}^2\text{-almost every } (i, j) \in \mathcal{S} \times I \setminus \mathcal{S}.$$

To prove this assertion, we consider for each  $i \in I$  the (possibly empty) sliced disconnection set

$$\mathcal{D}(i) := \left\{ j \in I \text{ s.t. } (i, j) \in \mathcal{D} \right\},$$

and observe that by Fubini's theorem, there exists a subset  $J \subset I$  of positive  $\mathcal{L}_{\perp I}^1$ -measure such that  $\mathcal{L}_{\perp I}^1(\mathcal{D}(i)) > 0$  for each  $i \in J$ . Given an element  $i \in J$ , let  $\mathcal{S}(i) \subset I$  be the largest strongly connected component containing  $i$ , which exists by Lemma 5.1, and notice that

$$\mathcal{L}_{\perp I}^1(\mathcal{S}(i)) > 0,$$

since  $\inf_{i \in I} \int_I a(i, j) dj > 0$  and  $a \in L^\infty(I \times I, [0, 1])$  is symmetric. In addition, one also has that  $\mathcal{L}_{\sqcup I}^1(\mathcal{S}(i) \cap \mathcal{D}(i)) = 0$  by definition (5.2) of the disconnection set, which then implies

$$\mathcal{L}_{\sqcup I}^1(I \setminus \mathcal{S}(i)) > 0.$$

Finally by the maximality of  $\mathcal{S}(i)$ , it necessarily holds that

$$a(k, l) = 0 \quad \text{for } \mathcal{L}_{\sqcup I \times I}^2\text{-almost every } (k, l) \in \mathcal{S}(i) \times I \setminus \mathcal{S}(i),$$

as we could otherwise build a strictly larger strongly connected set containing  $\mathcal{S}(i)$ . Armed with this construction, we consider the nontrivial element  $x \in \mathcal{C}^\perp \subset L^2(I, \mathbb{R}^d)$  defined by

$$x := \frac{1}{\mathcal{L}_{\sqcup I}^1(\mathcal{S}(i))} \mathbb{1}_{\mathcal{S}(i)} - \frac{1}{\mathcal{L}_{\sqcup I}^1(I \setminus \mathcal{S}(i))} \mathbb{1}_{I \setminus \mathcal{S}(i)}, \quad (5.3)$$

and observe that

$$\begin{aligned} \langle \mathbb{L}x, x \rangle_{L^2(I, \mathbb{R}^d)} &= \int_I \int_I a(k, l) \langle x(k), x(k) - x(l) \rangle dl dk \\ &= \int_{\mathcal{S}(i)} \int_I a(k, l) \langle x(k), x(k) - x(l) \rangle dl dk + \int_{I \setminus \mathcal{S}(i)} \int_I a(k, l) \langle x(k), x(k) - x(l) \rangle dl dk \\ &= \int_{\mathcal{S}(i)} \int_{\mathcal{S}(i)} a(k, l) \langle x(k), x(k) - x(l) \rangle dl dk + \int_{I \setminus \mathcal{S}(i)} \int_{I \setminus \mathcal{S}(i)} a(k, l) \langle x(k), x(k) - x(l) \rangle dl dk \\ &= 0, \end{aligned}$$

which contradicts the fact that  $\lambda_2(\mathbb{L}) > 0$ . Thus, we can conclude that the interaction topology defined by a symmetric adjacency operator  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is strongly connected as soon as  $\lambda_2(\mathbb{L}) > 0$ .

**Direct implication.** We now assume that the interaction topology defined by the symmetric adjacency operator  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is strongly connected in the sense of Definition 4.5. By Definition 2.5 and Proposition 2.6, the fact that  $\delta := \inf_{i \in I} \int_I a(i, j) dj > 0$  implies that  $\sigma_{\text{ess}}(\mathbb{L}) \subset [\delta, 1]$ . Recalling that the elements of  $\mathcal{C}$  belong to the kernel of  $\mathbb{L}$ , it follows from the min-max theorem that either  $\lambda_2(\mathbb{L}) > 0$ , or there exists a nontrivial element  $x \in \mathcal{C}^\perp$  for which  $\langle \mathbb{L}x, x \rangle_{L^2(I, \mathbb{R}^d)} = 0$ . By repeating the general argument of [16, Section 3.1], it can however be shown that the latter case is impossible.  $\square$

We end this section by showing how the proof strategy devised for the proof of Theorem 4.7 generalises to topologies representable as a disjoint union of strongly connected components.

**Theorem 5.3** (On the connectivity of directed graphons). *Let  $\mathcal{A} \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  be an adjacency operator defining a disjoint union of strongly components in the sense of Definition 4.5, and  $v \in L^\infty(I, \mathbb{R}_+^*)$  be the canonical eigenvector given by Theorem 4.7. Then, the interaction topology described by  $\mathcal{A}$  is strongly connected if and only if  $\lambda_2(\mathbb{L}_v) > 0$ , and in this case it holds that  $\inf_{i \in I} v(i) \geq \lambda_2(\mathbb{L}_v)$ .*

*Proof.* As already explained in the proof of Theorem 5.2 above, the direct implication of our claim can be deduced from the arguments of [16, Section 3.1] for general strongly connected graphons. Hence, we will focus on the converse one and assume that  $\lambda_2(\mathbb{L}_v) > 0$ . Observe first that the definition (4.18) of the algebraic connectivity can be rewritten as

$$\lambda_2(\mathbb{L}_v) = \inf_{x \in \mathcal{C}^\perp} \frac{\langle \mathbb{L}_v x, x \rangle_{L^2(I, \mathbb{R}^d)}}{\|x\|_{L^2(I, \mathbb{R}^d)}^2} = \inf_{x \in \mathcal{C}^\perp} \frac{\langle \mathbb{L}_v^{\text{sym}} x, x \rangle_{L^2(I, \mathbb{R}^d)}}{\|x\|_{L^2(I, \mathbb{R}^d)}^2},$$

where we denoted by  $\mathbb{L}_v^{\text{sym}} = \frac{1}{2}(\mathbb{L}_v + \mathbb{L}_v^*) \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  the symmetric part of the operator  $\mathbb{L}_v$ . The latter is a self-adjoint operator by construction, whose expression can be given explicitly as

$$\mathbb{L}_v^{\text{sym}} = \mathcal{M}_v \mathcal{M}_d - \frac{1}{2}(\mathcal{M}_v \mathcal{A} + \mathcal{A}^* \mathcal{M}_v),$$

where we recall that  $\mathcal{M}_d \in \mathcal{L}(L^2(I, \mathbb{R}^d))$  is the multiplication operator by the in-degree function defined as in (4.13). Moreover, it can be checked straightforwardly that the operator  $\mathcal{M}_v \mathcal{A} + \mathcal{A}^* \mathcal{M}_v$  is compact, and by reproducing the argument detailed in the proof of Theorem 5.2 above, one has

$$\sigma_{\text{ess}}(\mathbb{L}_v^{\text{sym}}) = \sigma_{\text{ess}}(\mathcal{M}_v \mathcal{M}_d) = \text{rg}(vd).$$

Upon noticing that  $\mathbb{L}_v^* x = 0$  for each  $x \in \mathcal{C}$ , we obtain by an application of the min-max theorem

$$0 < \lambda_2(\mathbb{L}_v) \leq \inf_{i \in I} v(i) \int_I a(i, j) dj. \quad (5.4)$$

Recalling that  $\|v\|_{L^\infty(I, \mathbb{R}_+^*)} \leq 1/\delta$  where  $\delta \in (0, 1]$  is such that (4.16) of Definition 4.5 holds, one has

$$\inf_{i \in I} \int_I a(i, j) dj \geq \delta \lambda_2(\mathbb{L}_v) > 0,$$

which yields Definition 4.5-(b). Using the fact that  $a \in L^\infty(I \times I, [0, 1])$ , we can also deduce from (5.4) that

$$\inf_{i \in I} v(i) \geq \lambda_2(\mathbb{L}_v) > 0.$$

There now remains to show that the conditions of Definition 4.5-(a) holds. To this end, we again assume by contradiction that the disconnection set  $\mathcal{D} \subset I \times I$  defined as in (5.2) has positive  $\mathcal{L}_{I \times I}^2$ -measure, and choose an element  $i \in I$  for which  $\mathcal{L}_{I \times I}^1(\mathcal{D}(i)) > 0$  where  $\mathcal{D}(i) = \{j \in I \text{ s.t. } (i, j) \in \mathcal{D}\}$ . Observe that by construction,  $i \in I_N$  for some  $N \geq 1$ . Thus by Lemma 5.1, the largest strongly connected component  $\mathcal{S}(i) \subset I$  containing  $i$  is necessarily given by  $\mathcal{S}(i) = I_N$ , and

$$\mathcal{L}_{I \times I}^1(\mathcal{S}(i)) = \mathcal{L}_{I \times I}^1(I_N) > 0.$$

Moreover, the fact that  $\mathcal{L}(\mathcal{S}(i) \cap \mathcal{D}(i)) = 0$  yields

$$\mathcal{L}_{I \times I}^1(I \setminus \mathcal{S}(i)) > 0,$$

and our standing assumption that the interaction topology is a disjoint union of strongly connected components further means

$$a(k, l) = 0 \quad \text{for } \mathcal{L}_{I \times I}^2\text{-almost every } (k, l) \in \mathcal{S}(i) \times I \setminus \mathcal{S}(i).$$

Then, by defining the non-trivial vector  $x \in \mathcal{C}^\perp$  as in (5.3), we can conclude that the condition  $\lambda_2(\mathbb{L}_v) = 0$  is violated, which ends the proof of our claim.  $\square$

**Remark 5.4** (On the lower-bound (4.47) imposed on the canonical eigenvectors in Theorem 4.13). *By applying Theorem 5.3 on every strongly connected set  $I_n$  while using (5.4), it can be checked that*

$$0 < \frac{1}{\mathcal{L}_{I \times I}^1(I_n)} \lambda_2(\mathbb{L}_{v_n}) \leq \inf_{i_n \in I_n} v_n(i_n),$$

where  $\mathbb{L}_{v_n} := \mathcal{M}_{v_n} \mathbb{L}_n \in \mathcal{L}(L^2(I_n, \mathbb{R}^d))$  denotes the restriction of the weighted graph-Laplacian to  $I_n$  for each  $n \geq 1$ . Thus, the lower-bound (4.47) holds whenever there exists a constant  $\nu > 0$  such that

$$\lambda_2(\mathbb{L}_{v_n}) \geq \nu \mathcal{L}_{I \times I}^1(I_n),$$

for each  $n \geq 1$ . This condition – that is in fact reminiscent of (4.16) – heuristically means that the total amount of interactions occurring within each of the countably many strongly connected components must be comparable to their size.

## 5.2 Equivalence between $L^2$ - and $L^\infty$ -consensus for strongly connected graphons

In this section, we prove an interesting and quite unexpected feature of consensus formation in graphon models, namely that convergence towards consensus in the  $L^2$ -norm topology implies convergence in the  $L^\infty$ -norm topology as soon as an average version of the lower-bound on the in-degree function imposed in Definition 4.5-(b) holds.

**Theorem 5.5** (Equivalence between  $L^2$ - and  $L^\infty$ -consensus formation). *Let  $x^0 \in L^\infty(I, \mathbb{R}^d)$  be such that  $\|x^0\|_{L^\infty(I, \mathbb{R}^d)} \leq R$  for some  $R > 0$ , assume that hypotheses **(GD)** hold, and denote by  $x(\cdot) \in \text{Lip}_{\text{loc}}(\mathbb{R}_+, L^2(I, \mathbb{R}^d))$  the corresponding solution of (2.8). Moreover, suppose that there exists a pair of coefficients  $(\tau, \mu) \in \mathbb{R}_+^* \times (0, 1]$  such that the persistence condition*

$$\frac{1}{\tau} \int_t^{t+\tau} \int_I a(s, i, j) dj ds \geq \mu, \quad (5.5)$$

holds for all times  $t \geq 0$  and  $\mathcal{L}^1_I$ -almost every  $i \in I$ . Then, one has that

$$\|x(t) - x^\infty\|_{L^\infty(I, \mathbb{R}^d)} \xrightarrow[t \rightarrow +\infty]{} 0 \quad \text{if and only if} \quad \|x(t) - x^\infty\|_{L^2(I, \mathbb{R}^d)} \xrightarrow[t \rightarrow +\infty]{} 0, \quad (5.6)$$

for any given element  $x^\infty \in \mathcal{C}(x^0)$ .

*Proof.* The direct implication in (5.6) is trivial since the  $L^2$ -norm is bounded from above by the  $L^\infty$ -norm on the compact set  $I := [0, 1]$ . Hence, we only need to prove the converse implication. To this end, we define for every  $\varepsilon > 0$  and  $\mathcal{L}^1$ -almost every  $t \geq 0$  the sets

$$I_\varepsilon(t) := \left\{ i \in I \text{ s.t. } |x(t, i) - x^\infty| \leq \varepsilon \right\},$$

and observe that by Chebyshev's inequality (see e.g. [1, Remark 1.18]), one has that

$$\mathcal{L}^1_I(I \setminus I_\varepsilon(t)) = \mathcal{L}^1_I\left(\left\{ i \in I \text{ s.t. } |x(t, i) - x^\infty| > \varepsilon \right\}\right) \leq \frac{1}{\varepsilon^2} \int_I |x(t, i) - x^\infty|^2 di,$$

for every  $\varepsilon > 0$ . Hence by the  $L^2$ -convergence posited in the right-hand side of (5.6), there exists a time horizon  $T_\varepsilon \geq 0$  such that  $\mathcal{L}^1_I(I \setminus I_\varepsilon(t)) \leq \varepsilon^2$  for  $\mathcal{L}^1$ -almost every  $t \geq T_\varepsilon$ . Recalling that  $x(\cdot)$  is a solution of the graphon dynamics (2.8), it also holds

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} |x(t, i) - x^\infty|^2 &= \int_I a(t, i, j) \phi(|x(t, i) - x(t, j)|) \langle x(t, i) - x^\infty, x(t, j) - x(t, i) \rangle dj \\ &= - \left( \int_I a(t, i, j) \phi(|x(t, i) - x(t, j)|) dj \right) |x(t, i) - x^\infty|^2 \\ &\quad + \int_I a(t, i, j) \phi(|x(t, i) - x(t, j)|) \langle x(t, i) - x^\infty, x(t, j) - x^\infty \rangle dj \\ &\leq -\gamma_R \left( \int_I a(t, i, j) dj \right) |x(t, i) - x^\infty|^2 \\ &\quad + c_\phi \left( \int_{I_\varepsilon(t)} |x(t, j) - x^\infty| dj + \int_{I \setminus I_\varepsilon(t)} |x(t, j) - x^\infty| dj \right) |x(t, i) - x^\infty| \\ &\leq -\gamma_R \left( \int_I a(t, i, j) dj \right) |x(t, i) - x^\infty|^2 + c_\phi \varepsilon R (1 + \varepsilon R), \end{aligned}$$

for  $\mathcal{L}^1$ -almost every  $t \geq T_\varepsilon$  and  $\mathcal{L}^1_I$ -almost every  $i \in I$ , where we recall that  $c_\phi = \sup_{r \in \mathbb{R}_+} \phi(r) < +\infty$  and  $\gamma_R = \min_{r \in [0, R]} \phi(2r)$ . Then by Grönwall's lemma, one further has

$$\begin{aligned} |x(t, i) - x^\infty|^2 &\leq |x(T_\varepsilon, i) - x^\infty|^2 \exp \left( -2\gamma_R \int_{T_\varepsilon}^t \int_I a(s, i, j) dj ds \right) \\ &\quad + c_\phi \varepsilon R (1 + \varepsilon R) \int_{T_\varepsilon}^t \exp \left( -2\gamma_R \int_s^t \int_I a(\sigma, i, j) dj d\sigma \right) ds, \end{aligned}$$

and by using the persistence condition (5.5) while repeating the computations of (3.17), we obtain

$$\begin{aligned}
|x(t, i) - x^\infty|^2 &\leq R^2 \exp\left(-2\gamma_R\mu(t - T_\varepsilon - \tau)\right) \\
&\quad + c_\phi\varepsilon R(1 + \varepsilon R) \exp(\gamma_R\mu\tau) \left(\frac{1 - \exp(-2\gamma_R\mu(t - T_\varepsilon))}{\gamma_R\mu}\right) \\
&\leq R^2 \exp\left(-2\gamma_R\mu(t - T_\varepsilon - \tau)\right) + \frac{\exp(\gamma_R\mu\tau)}{\gamma_R\mu} c_\phi\varepsilon R(1 + \varepsilon R),
\end{aligned} \tag{5.7}$$

for all times  $t \geq T_\varepsilon$  and  $\mathcal{L}_I^1$ -almost every  $i \in I$ . Thus given a parameter  $\delta > 0$ , we can choose  $\varepsilon > 0$  and then  $t \geq T_\varepsilon$  in such a way that

$$\frac{\exp(\gamma_R\mu\tau)}{\gamma_R\mu} c_\phi\varepsilon R(1 + \varepsilon R) \leq \frac{\delta^2}{2} \quad \text{and} \quad R^2 \exp\left(-2\gamma_R\mu(t - T_\varepsilon - \tau)\right) \leq \frac{\delta^2}{2},$$

which means that we can find a time horizon  $T_\delta \geq T_\varepsilon$  for which

$$|x(t, i) - x^\infty| \leq \delta \quad \text{for } \mathcal{L}_I^1\text{-almost every } i \in I,$$

whenever  $t \geq T_\delta$ . This is equivalent to stating that  $\|x(t) - x^\infty\|_{L^\infty(I, \mathbb{R}^d)}$  goes to 0 as  $t \rightarrow +\infty$ , which concludes the proof of our claim.  $\square$

**Remark 5.6** (On the persistence assumption (5.5)). *It can be checked that the lower-bound (5.5) on the average value of the in-degree function holds true e.g. when the algebraic connectivity is itself persistent – as in Theorem 4.11 or Theorem 4.13 –, or more generally if the interaction topology is algebraically persistent as in Theorem 4.8. This is again a consequence of the min-max theorem, as we recall that by (5.4), it holds that*

$$\begin{aligned}
\lambda_2(\mathbb{L}_v) &\leq \inf \left\{ \lambda \in \mathbb{R} \text{ s.t. } \lambda \in \sigma_{\text{ess}}(\mathbb{L}_v^{\text{sym}}) \right\} \\
&= \inf_{i \in I} \int_I v(i) a(i, j) dj \\
&\leq \|v\|_{L^\infty(I, \mathbb{R}^d)} \inf_{i \in I} \int_I a(i, j) dj,
\end{aligned}$$

for  $\mathcal{L}^1$ -almost every  $t \geq 0$ . Here,  $v \in L^\infty(I, \mathbb{R}_+^*)$  is the canonical eigenvector given either by Theorem 4.7 if the topology is a disjoint union of strongly connected components, or by  $v \equiv 1$  if the topology is either balanced or symmetric.

### 5.3 Numerical illustrations

In this last section, we provide illustrations of our results on a series of relevant examples. We start by showing positive exponential consensus results, first for a topology with persistent scrambling, then for a balanced topology with positive connectivity, and finally for a symmetric nonlinear topology that is algebraically persistent. We conclude by an example of balanced topology with null connectivity for which a non-exponential consensus only with respect to the  $L^2$ -norm seems to arise numerically.

**Exponential consensus under persistent scrambling.** In the following paragraphs, we start by exemplifying the consensus result established in Theorem 3.6 of Section 3. To this end, we fix a real parameters  $T > 0$  and an integer  $n \geq 1$ , and for any time  $t \geq 0$  we introduce the notation  $\mathfrak{t} := t - \lfloor t/T \rfloor \in [0, T)$ . We then define the interaction function  $a \in L^\infty(\mathbb{R}_+ \times I \times I, [0, 1])$  by

$$a(t, i, j) := \begin{cases} 1 - 2n \left| j - \frac{2n\mathfrak{t} + T}{2nT} \right| & \text{if } j \in \left[ \frac{\mathfrak{t}}{T}, \frac{\mathfrak{t}+1}{T} \right] \text{ and } \mathfrak{t} \in \left[ 0, \frac{n-1}{n}T \right), \\ 0 & \text{otherwise,} \end{cases} \tag{5.8}$$

for all times  $t \geq 0$  and every  $i, j \in I$ .



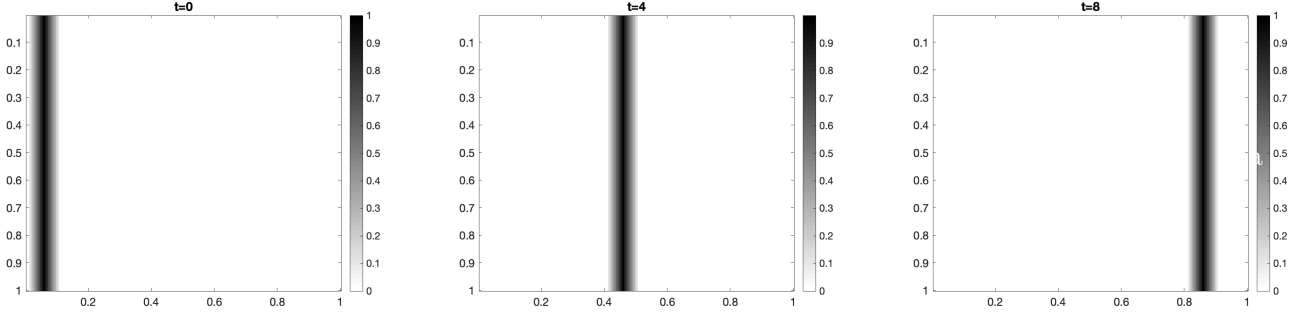


Figure 5.1: Representation of the interaction function  $(t, i, j) \in \mathbb{R}_+ \times I \times I \mapsto a(t, i, j) \in [0, 1]$  defined in (5.8) with  $T = n = 10$  at times  $t = 0$  (left),  $t = 4$  (center) and  $t = 8$  (right).

A first observation is that the value of  $a(t, i, j)$  defined in (5.8) is independent of  $i \in I$ , so that all the agents follow the same subset of elements at all times  $t \geq 0$ . As illustrated in Figure 5.1, this is a natural generalisation to the graphon setting of leader-follower dynamics (compare with the leftmost example in Figure 3.1), in which the leaders are modelled as a time-varying subset of agents with positive measure. Moreover, notice that the signal  $t \in \mathbb{R}_+ \mapsto a(t) \in L^\infty(I \times I, [0, 1])$  is  $T$ -periodic, and that the graphon is completely disconnected whenever  $\mathfrak{t} \in (\frac{n-1}{n}T, T)$ . In this context, we can explicitly compute the scrambling coefficients, since

$$\eta(\mathcal{A}(t)) = \inf_{i, j \in I} \int_I \min \{a(t, i, k), a(t, j, k)\} dk = \inf_{i \in I} \int_I a(t, i, k) dk = \frac{T}{2n},$$

whenever  $t \geq 0$  is such that  $\mathfrak{t} \in [0, \frac{n-1}{n}T)$ . On the other hand, it is clear that  $\eta(\mathcal{A}(t)) = 0$  if  $\mathfrak{t} \in [\frac{n-1}{n}T, T)$ . Thus for every parameter  $\tau > T/n$ , one can check that for all times  $t \geq 0$ , it holds

$$\frac{1}{\tau} \int_t^{t+\tau} \eta(\mathcal{A}(s)) ds \geq \frac{1}{\tau} \int_{(n-1)T/n}^{T+(\tau-T/n)} \eta(\mathcal{A}(s)) ds \geq \frac{T(n\tau - T)}{2\tau n^2},$$

so that the persistence condition (3.14) of Theorem 3.6 is satisfied. Therefore, we expect that the system will converge exponentially to consensus with respect to the  $L^\infty$ -norm, a fact that is illustrated by the numerical simulations displayed in Figure 5.2 below. Therein, we fix the parameters  $T = n = 10$  and start from the initial configuration  $x^0 : i \in I \mapsto \sin^2(4i) \in [0, 1]$ .

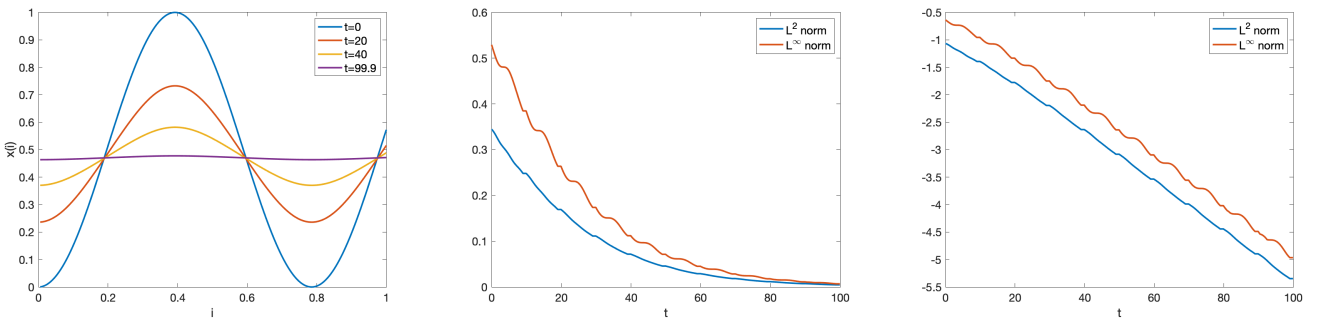


Figure 5.2: Snapshots of the solution  $i \in I \mapsto x(t, i) \in [0, 1]$  generated by the communication weights (5.8) at different instants (left) along with the time-evolution of the  $L^2$ - and  $L^\infty$ -distance to the consensus point in natural scale (center) and log scale (right).

**Exponential consensus for a balanced topology with positive connectivity** This time, we consider a stationary balanced interaction topology which is defined as follows. Let  $\xi \in C^0([0, 1])$  be the function given by

$$\xi(s) := (1 - 4s)\mathbb{1}_{[0, \frac{1}{4}]}(s) \quad \text{for each } s \in [0, 1],$$

and, by a slight abuse of notation, denote again by  $\xi \in C^0(\mathbb{R}, [0, 1])$  its 1-periodisation over the real line. Then, we consider the constant-in-time interaction kernel  $a \in L^\infty(I \times I, [0, 1])$  defined by

$$a(i, j) := \xi(i - j), \quad (5.9)$$

for every  $i, j \in I$ , which is illustrated in Figure 5.3. This particular choice of interaction function encodes the notion of directed cycle at the level of graphon models (draw a comparison with the leftmost topology in Figure 4.1), where the neighbours are represented by small clusters of agents with positive measure.

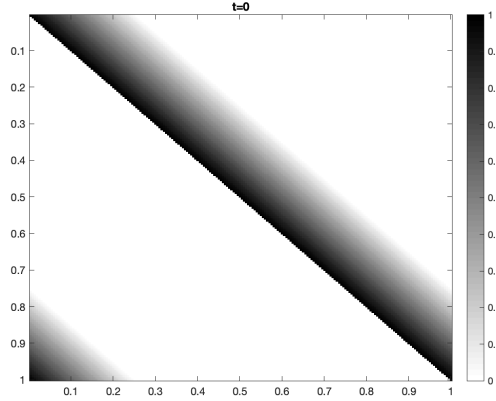


Figure 5.3: Representation of the interaction function  $(i, j) \in I \times I \mapsto a(i, j) \in [0, 1]$  defined in (5.9).

It can be readily verified that the interaction kernel  $a \in L^\infty(I \times I, [0, 1])$  defined as in (5.9) generates a balanced topology, since

$$\begin{aligned} \int_I a(i, j) dj &= \int_I \xi(i - j) dj = \int_0^i \xi(i - j) dj + \int_i^1 \xi(i - j + 1) dj \\ &= \int_0^i \xi(s) ds + \int_i^1 \xi(s') ds' = \int_0^1 \xi(s) ds = \int_I a(j, i) dj, \end{aligned} \quad (5.10)$$

where we used the fact that  $\xi(\cdot)$  is 1-periodic and applied the changes of coordinates  $s = i - j$  and  $s' = 1 + i - j$  in the first and second integral respectively. We now prove that the topology is strongly connected in the sense of Definition 4.5. It can be verified straightforwardly by using (5.10) that

$$\inf_{i \in I} \int_I a(i, j) dj = \int_0^1 \xi(s) ds = \int_0^{1/4} (1 - 4s) ds = \frac{1}{8},$$

so that the condition of Definition 4.5-(b) holds. Moreover, for every pair  $i, j \in I$  with  $i < j$ , define the finite sequence of indices  $(l_k)_{1 \leq k \leq 8}$  by

$$l_1 = i \quad \text{and} \quad l_{k+1} = \min \left\{ l_k + \frac{1}{8}, j \right\} \quad \text{for } k \in \{1, \dots, 7\},$$

and observe that  $j = l_{K_j}$  for some  $K_j \in \{1, \dots, 7\}$  with  $a(l_k, l_{k+1}) \geq \frac{1}{2}$  for each  $k \in \{1, \dots, K_j - 1\}$ . Therefore, the conditions of Definition 4.5-(a) are satisfied, and the interaction topology is strongly connected. Hence by Theorem 5.3, it holds that  $\lambda_2(\mathbb{L}) > 0$  and the system should converge exponentially towards consensus in the  $L^2$ -norm. In addition, because the sufficient condition of Theorem 5.5 is satisfied, the convergence should also occur in the  $L^\infty$ -norm topology. The simulations made with the initial datum  $x^0 : i \in I \mapsto \sin^2(4i) \in [0, 1]$  and displayed in Figure 5.4 confirm these assertions. In addition, let it be noted that the convergence with respect to the  $L^\infty$ -norm also seems to be exponential, which could suggest that a stronger variant of Theorem 5.5 stating the equivalence between exponential convergences towards consensus may in fact be true.

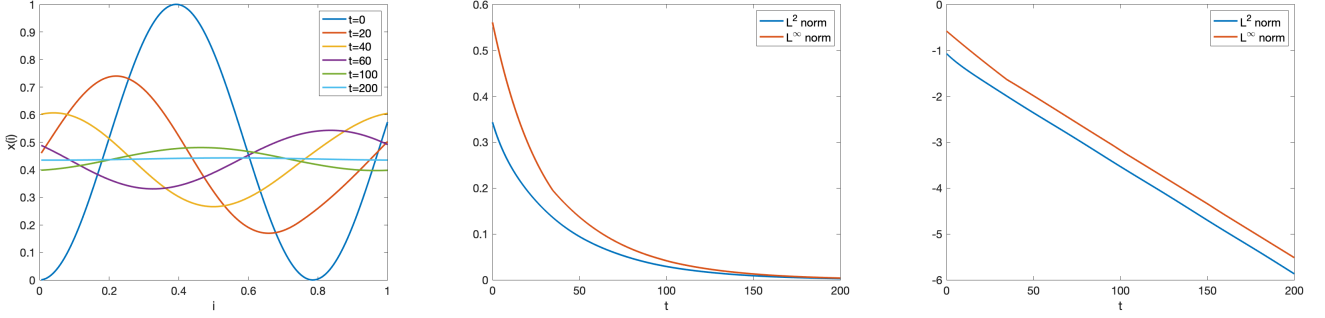


Figure 5.4: Snapshots of the solution  $i \in I \mapsto x(t, i) \in [0, 1]$  generated by the communication weights (5.9) at different instants (left) along with the time-evolution of the  $L^2$ - and  $L^\infty$ -distance to the consensus point in natural scale (center) and log scale (right).

### Exponential consensus for an algebraically persistent symmetric nonlinear topology.

Next, we turn our attention to a nonlinear symmetric and time-dependent interaction kernel which defines disconnected topologies at all times  $t \geq 0$ , but which is algebraically persistent in the sense defined in Theorem 4.8. Given a real parameter  $T > 0$  and an integer  $n \geq 1$ , we introduce the notation  $\mathfrak{t} := t - \lfloor t/T \rfloor \in [0, T)$ . We then define the interaction function  $a \in L^\infty(\mathbb{R}_+ \times I \times I, [0, 1])$  by

$$a(t, i, j) := \begin{cases} 1 & \text{if } i, j \in [\frac{\mathfrak{t}}{T}, \frac{\mathfrak{t}}{T} + \frac{1}{n}] \text{ and } \mathfrak{t} \in [0, (\frac{n-1}{n})T), \\ 1 & \text{if } i, j \in [\frac{\mathfrak{t}}{T}, 1] \cup [0, \frac{\mathfrak{t}}{T} + \frac{1}{n} - 1] \text{ and } \mathfrak{t} \in [(\frac{n-1}{n})T, T), \\ 0 & \text{otherwise,} \end{cases} \quad (5.11)$$

for all times  $t \geq 0$  and every  $i, j \in I$ . The latter is symmetric for all times  $t \geq 0$  by construction, and is illustrated in Figure 5.5. We also consider the nonlinear Cucker-Smale type kernel given by

$$\phi(r) := \frac{1}{(1+r)^2},$$

for each  $r \geq 0$ , and consider the solution  $x(\cdot)$  of (4.9) with initial condition  $x^0 : i \in I \mapsto \sin^2(4i) \in [0, 1]$ .

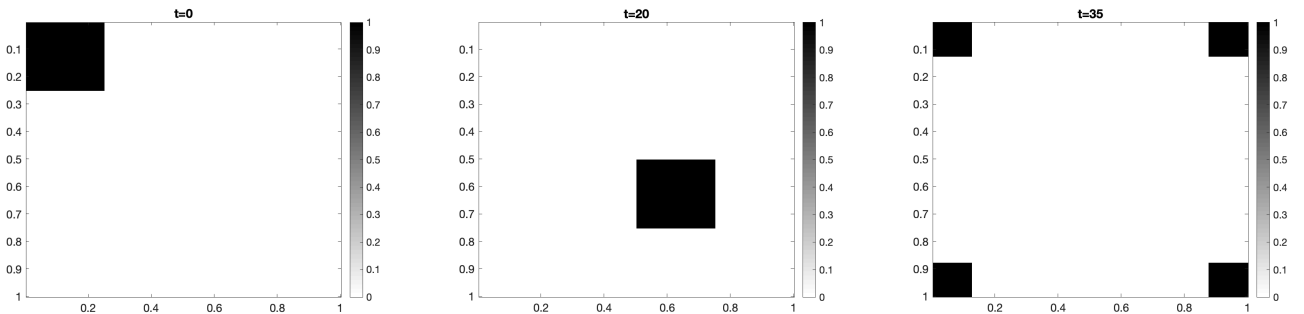


Figure 5.5: Representation of the interaction function  $(t, i, j) \in \mathbb{R}_+ \times I \times I \mapsto a(t, i, j) \in [0, 1]$  defined in (5.11) with  $T = 40$  and  $n = 4$  at times  $t = 0$  (left),  $t = 20$  (center) and  $t = 35$  (right).

It is quite clear from its definition in (5.11) that in our example, the interaction topology is disconnected at each instant. However, the average of the interaction function over any time window of length  $T$  defines the counterpart for graphon models of an undirected cycle, which is illustrated in Figure 5.6 (to be compared with the central picture in Figure 3.1). Indeed, for all  $i, j \in I$  satisfying

$$j \in \left[ \max \left\{ 0, i - \frac{1}{n} \right\}, \min \left\{ 1, i + \frac{1}{n} \right\} \right],$$

it can be verified that

$$\frac{1}{T} \int_t^{t+T} \int_I a(s, i, j) dj ds = \frac{1}{T} \int_0^T \int_I a(s, i, j) dj ds \geq \frac{1}{n^2},$$

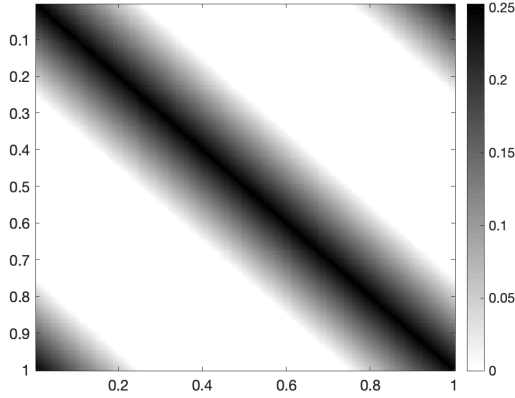


Figure 5.6: Representation of the average over a time-window of length  $T = 10$  of the interaction function  $(t, i, j) \in I \times I \mapsto a(t, i, j) \in [0, 1]$  defined in (5.11).

for all times  $t \geq 0$ , where we used the fact that signal  $t \in \mathbb{R}_+ \mapsto a(t) \in L^\infty(I \times I, [0, 1])$  is  $T$ -periodic. Therefore by Theorem 5.2, the algebraic connectivity of the averaged interaction topology is positive, and the hypotheses of Theorem 4.8 are satisfied. Thus, as confirmed by the results displayed in Figure 5.7, we expect that the system exponentially converges to consensus in the  $L^2$ -norm topology. In addition, it can be readily verified that the in-degree function satisfies the persistence condition of Theorem 5.5, so that convergence towards consensus also holds with respect to the  $L^\infty$ -norm.

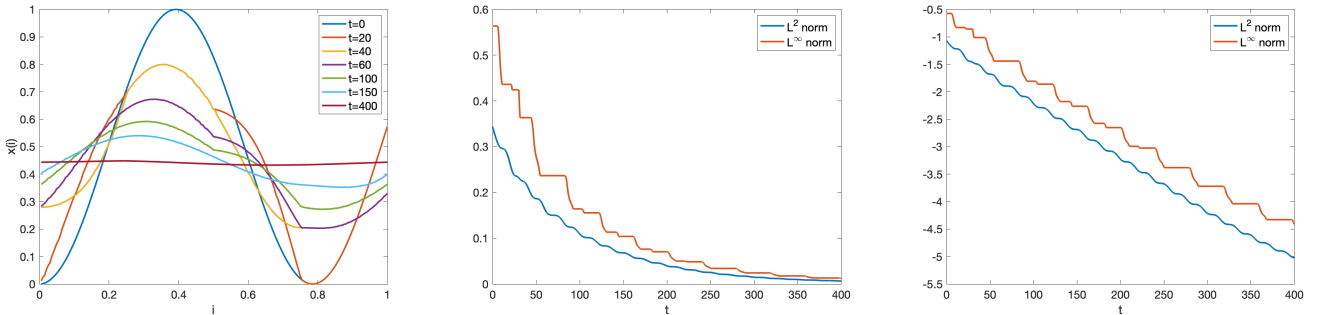


Figure 5.7: Snapshots of the solution  $i \in I \mapsto x(t, i) \in [0, 1]$  generated by the communication weights (5.11) at different instants (left) along with the time-evolution of the  $L^2$ - and  $L^\infty$ -distance to the consensus point in natural scale (center) and log scale (right).

**Non-exponential consensus for a symmetric topology with null connectivity.** In this last paragraph, we provide an interesting limit example of the theory developed in this article, which highlights several of our results. To this end, we consider the stationary interaction kernel defined by

$$a(i, j) := \begin{cases} 1 & \text{if either } i \in [0, \frac{1}{2}] \text{ and } \frac{i}{2} \leq j \leq 2i \text{ or } i, j \in [\frac{1}{2}, 1], \\ 0 & \text{otherwise,} \end{cases} \quad (5.12)$$

for every  $i, j \in I$ , and which is illustrated in Figure 5.8.

It is clear from its definition that  $a \in L^\infty(I \times I, [0, 1])$  defines a symmetric topology, but one can readily check that  $\inf_{i \in I} \int_I a(i, j) dj = 0$ , which implies that the interaction graphon is not strongly connected in the sense of Definition 4.5. By Theorem 5.2, this implies in particular that the corresponding algebraic connectivity satisfies  $\lambda_2(\mathbb{L}) = 0$ , which means that even though the convergence to consensus in the  $L^2$ -norm may occur, it may not be exponential. Furthermore, the sufficient condition of Theorem 5.5 ensuring that  $L^2$ - and  $L^\infty$ -consensus formation are equivalent is also violated,

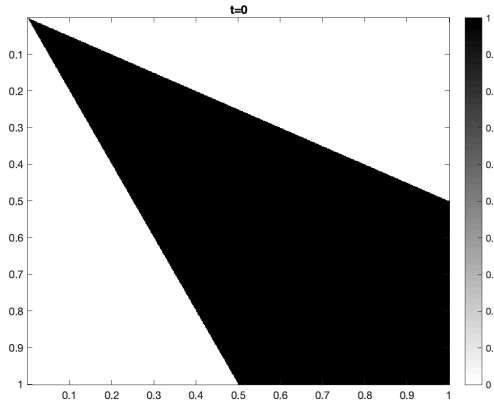


Figure 5.8: Representation of the interaction function  $(i, j) \in I \times I \mapsto a(i, j) \in [0, 1]$  defined in (5.12).

which suggests that solutions of the graphon dynamics may not converge to consensus in the  $L^\infty$ -norm topology. This intuition is supported by the fact that, in this precise example, the agent with label  $i = 0$  is not connected to any other agent in the system, and should not move.

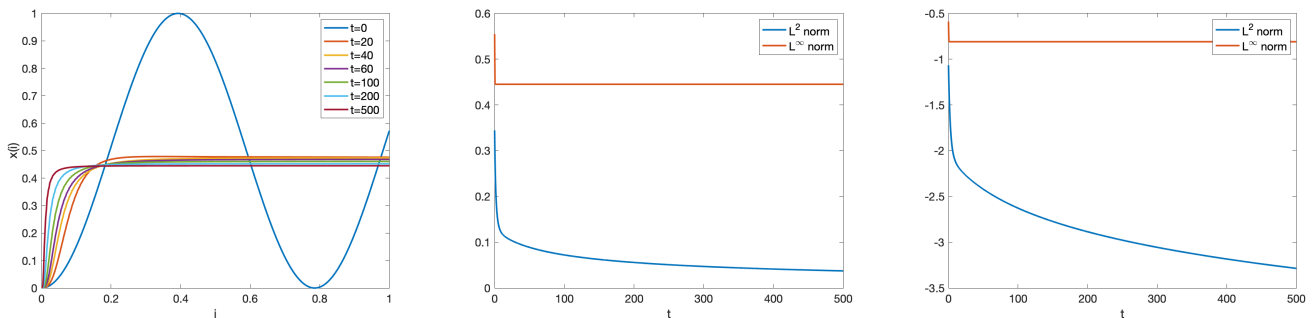


Figure 5.9: Snapshots of the solution  $i \in I \mapsto x(t, i) \in [0, 1]$  generated by the communication weights (5.12) at different instants (left) along with the time-evolution of the  $L^2$ - and  $L^\infty$ -distance to the consensus point in natural scale (center) and log scale (right).

The plots displayed in Figure 5.9 are coherent with what was expected, in the sense that the  $L^\infty$ -norm remains bounded from below by a constant, while a non-exponential  $L^2$ -consensus formation seems to arise. To understand why the latter appears – at least numerically – even though  $\lambda_2(\mathbb{L}) = 0$ , recall that the semi-discretisation of the graphon dynamics is given for any chosen integer  $N \geq 1$  by

$$\dot{x}_i(t) = \frac{1}{N} \sum_{j=1}^N a_{ij}(x_j(t) - x_i(t)), \quad x_i(0) = x_i^0, \quad (5.13)$$

for all times  $t \geq 0$  and every  $i \in \{1, \dots, N\}$ , where

$$a_{ij} := N^2 \int_{(2i-1)/2N}^{(2i+1)/2N} \int_{(2j-1)/2N}^{(2j+1)/2N} a(k, l) dl dk \quad \text{and} \quad x_i^0 := \int_{(2i-1)/2N}^{(2i+1)/2N} x^0(k) dk, \quad (5.14)$$

for every  $i, j \in \{1, \dots, N\}$ . The key point here is to remark that the adjacency matrices  $\mathbf{A}_N := (a_{ij})_{1 \leq i, j \leq N} \in [0, 1]^N$  – that are represented for different values of  $N$  in Figure 5.10 – define strongly connected interaction topologies with  $\lambda_2(\mathbf{L}_N) > 0$  and  $\min_{i \in \{1, \dots, N\}} \frac{1}{N} \sum_{j=1}^N a_{ij} > 0$ . Therefore, at the discrete level, we always observe the formation of  $\ell^2$ - and  $\ell^\infty$ -consensus as illustrated in Figure 5.11.

The fact that the exponential convergence to consensus of the sequence of approximating microscopic problems is not a stable property as  $N \rightarrow +\infty$  can be further underpinned by the following

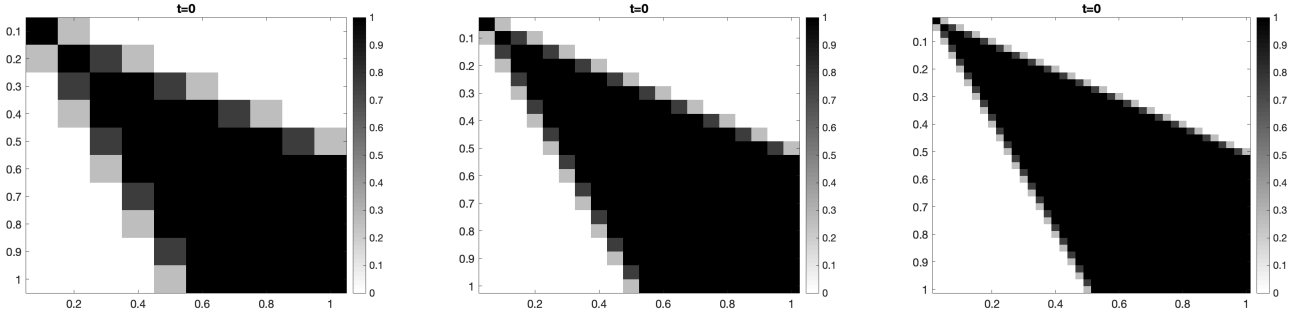


Figure 5.10: Representation of the interaction matrices  $\mathbf{A}_N := (a_{ij})_{1 \leq i, j \leq N}$  whose coefficients are defined by (5.14) for  $N = 10$  (left),  $N = 20$  (center) and  $N = 40$  (right).

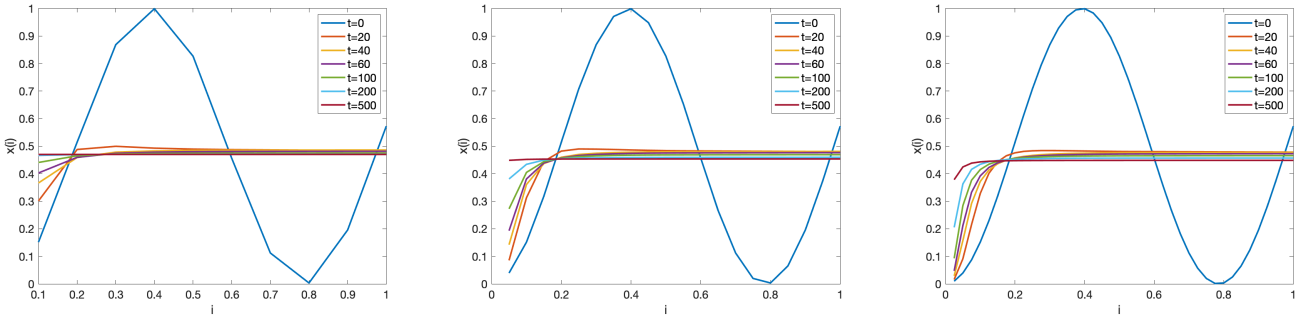


Figure 5.11: Snapshots of the solution  $i \in I \mapsto x(t, i) \in [0, 1]$  generated by the communication weights (5.12) at different instants for  $N = 10$  (left),  $N = 20$  (center) and  $N = 40$  (right).

observation. As amply discussed in a more general context in [77], only the discrete part of the spectrum of a graph-Laplacian operator can be faithfully described by the limits of eigenvalues of discrete approximating graph-Laplacian matrices. In our example, the essential spectrum of the operator is  $\sigma_{\text{ess}}(\mathbb{L}) = \text{rg}(d) = [0, 1]$ , and the sequence  $(\lambda_2(\mathbf{L}_N))_{N \geq 1}$  of algebraic connectivities of the discrete models goes to 0 as  $N \rightarrow +\infty$ . This induces a depreciation of the convergence rates as  $N$  gets larger, which is illustrated for both norms in Figure 5.12.

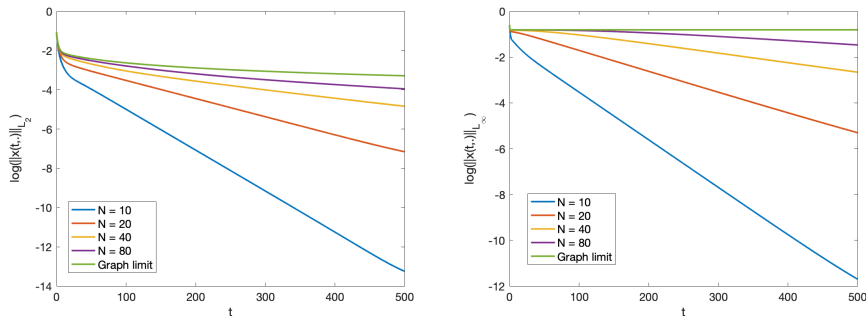


Figure 5.12: Evolution of the long-time convergence rates to consensus of solutions the discrete systems (5.13) with respect to the  $L^2$ -norm (left) and the  $L^\infty$ -norm (right) as  $N$  increases.

**Acknowledgement.** The authors wish to thank Christophe Hazard and Emmanuel Trélat for their useful insights and bibliographical suggestions.

## References

- [1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variations and Free Discontinuity Problems*. Oxford Mathematical Monographs, 2000.
- [2] J.-P. Aubin and A. Cellina. *Differential Inclusions*. Springer-Verlag, 1984.
- [3] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Modern Birkhäuser Classics. Birkhäuser Basel, 1990.
- [4] N. Ayi and N. Pouradier Duteil. Mean-Field and Graphs Limits for Collective Dynamics Models with Time-Varying Weights. *Journal of Differential Equations*, 299:65–110, 2021.
- [5] M. Ballerini, N. Cabibbo, R. Candelier, et al. Interaction Ruling Animal Collective Behavior Depends on Topological Rather than Metric Distance: Evidence from a Field Study. *Proceedings of the National Academy of Sciences*, 105(4):1232–1237, 2008.
- [6] R.W. Beard and W. Ren. *Distributed Consensus in Multi-Vehicle Cooperative Control*. Springer-Verlag, 2008.
- [7] N. Bellomo, M. A. Herrero, and A. Tosin. On the Dynamics of Social Conflicts: Looking for the Black Swan. *Kinetic & Related Models*, 6(3):459–479, 2013.
- [8] H. Berliocchi and J.-M. Lasry. Intégrandes Normales et Mesures Paramétrées en Calcul des Variations. *Bulletins de la Société Mathématique de France*, 101:129–184, 1973.
- [9] A.L. Bertozzi, T. Kolokolnikov, H. Sun, D. Uminsky, and J. Von Brecht. Ring Patterns and their Bifurcations in a Nonlocal Model of Biological Swarms. *Communications in Mathematical Sciences*, 13(4):2015, 2015.
- [10] U. Biccari, K. Dongnam, and E. Zuazua. Dynamics and Control for Multi-Agent Networked Systems: A Finite-Difference Approach. *Mathematical Models and Methods in Applied Sciences*, 29(4):755–790, 2019.
- [11] B. Birnir. An ODE Model of the Motion of Pelagic Fish. *Journal of Statistical Physics*, 128(1):535–568, 2007.
- [12] V.D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. On Krause’s Multi-Agent Consensus Model with State-Dependent Connectivity (Extended version). *IEEE Transactions on Automatic Control*, 54(11):2586–2597, 2009.
- [13] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis. Convergence in Multiagent Coordination, Consensus, and Flocking. In *Proceedings of the 2005 44th IEEE Conference on Decision and Control (CDC)*, pages 2996–3000, 2005.
- [14] B. Bonnet and É. Flayac. Consensus and Flocking under Communication Failures for a Class of Cucker-Smale Systems. *System and Control Letters*, 152:104930, 10, 2021.
- [15] B. Bonnet and H. Frankowska. Differential Inclusions in Wasserstein Spaces: The Cauchy-Lipschitz Framework. *Journal of Differential Equations*, 271:594–637, 2021.
- [16] L. Boudin, F. Salvarini, and E. Trélat. Exponential Convergence Towards Consensus for Non-Symmetric Linear First-Order Systems in Finite and Infinite Dimensions. *SIAM Journal on Mathematical Analysis*, 54(3):2727–2752, 2022.
- [17] F. Bullo, J. Cortés, and S. Martinez. *Distributed Control of Robotic Networks*. Applied Mathematics. Princeton University Press, 2009.
- [18] P.E. Caines and M.Y. Huang. Graphon Mean Field Games and the GMFG Equations:  $\varepsilon$ -Nash Equilibria. In *2019 IEEE 58th Conference on Decision and Control (CDC), Nice, France*, pages 286–292, 2019.
- [19] M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat. Sparse Stabilization and Optimal Control of the Cucker-Smale Model. *Mathematical Control and Related Fields*, 3(4):447–466, 2013.
- [20] M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat. Sparse Stabilization and Control of Alignment Models. *Mathematical Models and Methods in Applied Sciences*, 25 (3):521–564, 2015.
- [21] R. Carmona, C.V. Cooney, D.B. Graves, and M. Laurière. Stochastic Graphon Games: I. The Static Case. *Mathematics of Operations Research*, 47:750–778, 2021.
- [22] J.A. Carrillo, M. Fornasier, J. Rosado, and G. Toscani. Asymptotic Flocking for the Kinetic Cucker-Smale Model. *SIAM Journal on Mathematical Analysis*, 42(1):218–236, 2010.
- [23] J.A. Carrillo, D. Kalise, F. Rossi, and E. Trélat. Controlling Swarming Models Towards Flocks and Mills. *arxiv preprint arXiv:2103.07304v3*, 2021.
- [24] Y.-P. Choi, J.A. Carrillo, and M. Hauray. The Derivation of Swarming Models: Mean-Field Limit and Wasserstein Distances. In *Collective Dynamics from Bacteria to Crowds, CISM International Centre for Mechanical Sciences*, volume 553, pages 1–46. Springer, 2014.
- [25] Y.-L. Chuang, M.R. D’Orsogna, D. Marthaler, A.L. Bertozzi, and L.S. Chayes. State Transitions and the Continuum Limit for a 2D Interacting, Self-Propelled Particle System. *Physica D: Nonlinear Phenomena*, 232(1):33–47, 2007.
- [26] E. Cristiani, B. Piccoli, and A. Tosin. *Multiscale Modeling of Pedestrian Dynamics*, volume 12. Springer, 2014.
- [27] F. Cucker and S. Smale. Emergent Behavior in Flocks. *IEEE Transactions on Automatic Control*, 52(5):852–862, 2007.
- [28] F. Cucker and S. Smale. On the Mathematics of Emergence. *Japanese Journal of Mathematics*, 2(1):197–227, 2007.

- [29] F. Dalmao and E. Mordecki. Cucker-Smale Flocking under Hierarchical Leadership and Random Interactions. *SIAM J. Appl. Math.*, 71(4):1307–1316, 2011.
- [30] J.M. Danskin. *The Theory of Max-Min and its Application to Weapons Allocation Problems*, volume 5 of *Ökonometrie und Unternehmensforschung Econometrics and Operations Research*. Springer-Verlag Berlin Heidelberg, 1967.
- [31] J. Diestel and J.J.Jr Uhl. *Vector Measures*. Number 15 in Mathematical Surveys. American Mathematical Society, 1977.
- [32] K.-J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, 2001.
- [33] L.C. Evans and R.F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, 1992.
- [34] M. Fiedler. Algebraic Connectivity of Graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973.
- [35] A.F. Filippov. *Differential Equations with Discontinuous Righthand Sides: Control Systems*, volume 18 of *Springer Science and Business Media*. Springer, 2013.
- [36] G. Flierl, D. Grünbaum, S. Levin, and D. Olson. From Individuals to Aggregations: the Interplay between Behavior and Physics. *Journal of Theoretical Biology*, 196:397–454, 1999.
- [37] S. Gao and P.E. Caines. Graphon Control of Large-Scale Networks of Linear Systems. *IEEE Transaction on Automatic Control*, 65(10):4090–4105, 2019.
- [38] S.-Y. Ha and J.G. Liu. A Simple Proof of the Cucker-Smale Flocking Dynamics and Mean-Field Limit. *Comm. Math. Sci.*, 7(2):297–325, 2009.
- [39] S.-Y. Ha and E. Tadmor. From Particle to Kinetic and Hydrodynamic Descriptions of Flocking. *Kinetic and Related Models*, 1:415–435, 2008.
- [40] V. Hardt and E. Wagenführer. Spectral Properties of a Multiplication Operator. *Mathematische Nachrichten*, 178:135–156, 1996.
- [41] R. Hegselmann and U. Krause. Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.
- [42] S. Heinrich and P. Mankiewicz. Applications of Ultrapowers to the Uniform and Lipschitz Classification of Banach Spaces. *Studia Mathematica*, 73:225–251, 1982.
- [43] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Texts. Cambridge, 1985.
- [44] P.-E. Jabin and S. Motsch. Clustering and Asymptotic Behavior in Opinion Formation. *Journal of Differential Equations*, 257(11):4165–4187, 2014.
- [45] A. Jadbabaie, J. Lin, and A.S. Morse. Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules. *IEEE Transactions in Automatic Control*, 48(6):988 –1001, 2003.
- [46] U. Krause. A Discrete Nonlinear and Non-Autonomous Model of Consensus Formation. volume 2000, pages 227–236. Gordon and Breach, Amsterdam, 2000.
- [47] B.E. Lee. Consensus and Voting on Large Graphs: An Application of Graph Limit Theory. *Discrete and Continuous Dynamical Systems - Series A*, 38(4):1719–1744, 2018.
- [48] D. Liberzon. *Switching in Systems and Control*. Systems and Control : Foundations and Applications. Birkhäuser Boston, 2003.
- [49] L Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012.
- [50] L. Lovász and B. Szegedy. Limits of Dense Graph Sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, 2006.
- [51] S. Manfredi and D. Angeli. A Criterion for Exponential Consensus of Time-Varying Non-Monotone Nonlinear Networks. *IEEE Transactions on Automatic Control*, 62(5):2483 – 2489, 2016.
- [52] S. Martin, A. Girard, A. Fazeli, and A. Jadbabaie. Multi-Agent Flocking under General Communication Rule. *IEEE Transactions on Control and Network Systems*, 1(2):155–166, 2014.
- [53] F. Mazenc and M. Malisoff. *Construction of Strict Lyapunov Functions*. Springer-Verlag, 2009.
- [54] S. McQuade, B. Piccoli, and N. Pouradier Duteil. Social Dynamics Models with Time-Varying Influence. *Mathematical Models and Methods in Applied Sciences*, 29(04):681–716, 2019.
- [55] G.S. Medvedev. The Nonlinear Heat Equation on Dense Graphs and Graph Limits. *SIAM Journal on Mathematical Analysis*, 46:2743–2766, 2014.
- [56] G.S. Medvedev. The Nonlinear Heat Equation on W-Random Graphs. *Archive for Rational Mechanics and Analysis*, 212:pages 781–803, 2014.
- [57] M. Mesbahi and M. Egerstedt. *Graph Theoretic Methods in Multiagent Networks*. 2010.
- [58] B. Mohar. The laplacian spectrum of graphs. In *Graph Theory, Combinatorics, and Applications*, volume 2, pages 871–898. Wiley, 1991.



- [59] L. Moreau. Stability of Multiagent Systems with Time-Dependent Communication Links. *IEEE Transactions on Automatic Control*, 50(2):169–182, 2005.
- [60] S. Motsch and E. Tadmor. A New Model for Self-Organized Dynamics and its Flocking Behavior. *Journal of Statistical Physics*, 144(923):923–947, 2011.
- [61] S. Motsch and E. Tadmor. Heterophilious Dynamics Enhances Consensus. *SIAM Review*, 56(4):577–621, 2014.
- [62] A. Muntean, J. Rademacher, and A. Zagaris. *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*. Springer, 2016.
- [63] K-S. Narendra and A-M. Annaswami. *Stable Adaptive Systems*. Dover Books on Electrical Engineering. Dover, 1989.
- [64] R. Olfati-Saber and R.M. Murray. Consensus Problems in Networks of Agents with Switching Topology and Time-Delays. *IEEE Transaction on Automatic Control*, 49(9):1520–1533, 2004.
- [65] B. Piccoli and F. Rossi. Transport Equation with Nonlocal Velocity in Wasserstein Spaces : Convergence of Numerical Schemes. *Acta Applicandae Mathematicae*, 124(1):73–105, 2013.
- [66] B. Piccoli and F. Rossi. Measure Theoretic Models for Crowd Dynamics. In *Crowd Dynamics Volume 1 - Theory, Models, and Safety Problems*, pages 137–165. Birkhauser, 2018.
- [67] B. Piccoli, F. Rossi, and E. Trélat. Control to Flocking of the Kinetic Cucker-Smale model. *SIAM Journal on Mathematical Analysis*, 47(6):4685–4719, 2015.
- [68] M. Reed and B. Simon. *Analysis of Operators*, volume 4 of *Methods of Modern Mathematical Physics*. Academic Press, 1978.
- [69] M. Reed and B. Simon. *Functional Analysis*, volume 1 of *Methods of Modern Mathematical Physics*. Academic Press, 1981.
- [70] L. Ruiz, L.F.O. Chamon, and A. Ribeiro. Graphon Neural Networks and the Transferability of Graph Neural Networks. In *34th Conference on Neural Information Processing Systems (NeurIPS 2020)*, page 1702–1712, 2020.
- [71] E. Seneta. Coefficients of Ergodicity: Structure and Applications. *Advances in Applied Probability*, 11(3):576–590, 1979.
- [72] J. Shen. Cucker–Smale Flocking under Hierarchical Leadership. *SIAM Journal on Applied Mathematics*, 68(3):694–719, 2008.
- [73] A. Tahbaz-Salehi and A. Jadbabaie. A Necessary and Sufficient Condition for Consensus Over Random Networks. *IEEE Transaction on Automatic Control*, 53(3):791–795, 2008.
- [74] H.G. Tanner, A. Jadbabaie, and G.J Pappas. Flocking in fixed and switching networks. *IEEE Transactions on Automatic Control*, 52(5):863–868, 2007.
- [75] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel Type of Phase Transition in a System of Self-Driven Particles. *Physical Review Letters*, 75(6):1226–1229, 1995.
- [76] R. Vizuete, F. Garin, and P. Frasca. The Laplacian Spectrum of Large Graphs Sampled from Graphons. *IEEE Transactions on Network Science and Engineering*, 8(2):1711–1721, 2021.
- [77] U. von Luxburg, M. Belkin, and O. Bousquet. Consistency of Spectral Clustering. *The Annals of Statistics*, 36(2):555–586, 2008.
- [78] D.B. West. *Introduction to Graph Theory*. Prentice Hall, 2001.
- [79] C.W. Wu. Algebraic Connectivity of Directed Graphs. *Linear and Multilinear Algebra*, 53(3):203–223, 2005.
- [80] C.W. Wu. On Bounds of Extremal Eigenvalues of Irreducible and m-Reducible Matrices. *Linear Algebra and its Applications*, 402:29–45, 2005.
- [81] C.W. Wu. On the Rayleigh-Ritz Ratio of a Generalized Laplacian Matrix of Directed Graphs. *Linear Algebra and its Applications*, 402:207–227, 2005.