Asymptotic consensus with transmission and reaction delay: an overview

Jan Haskovec*

Abstract

The aim of this paper is to provide a systematic overview of results on asymptotic consensus for the Hegselmann-Krause-type model with delay and discuss the corresponding analytical tools. We explain that two types (sources) of delay – transmission and reaction – are justifiable from the modeling point of view. We consider both classical and normalized communication weights. Studying a toy model with two agents only, we develop an intuitive insight into what type of dynamics we can expect from the systems. In particular, we stress that with transmission-type delay, asymptotic consensus can be reached with any length of the delay (i.e., without smallness assumptions). In contrast, the systems with reaction-type delay can only reach asymptotic consensus if the delay is sufficiently small.

We formulate four theorems that establish asymptotic consensus in the following scenarios: (1) transmission-type delay with classical communication weights, (2) transmission-type delay with normalized communication weights, (3) reaction-type delay with symmetric communication weights, (4) reaction-type delay with non-symmetric communication weights. We explain how the methods of proof depend on the particular scenario: direct estimates for (1), convexity arguments for (2), Lyapunov functional for (3) and generalized Gronwall-Halanay inequality for (4).

Keywords: Hegselmann-Krause model, time delay, asymptotic consensus, long-time behavior.

1 Introduction

The Hegselmann-Krause model [22] of opinion dynamics is a generic model of self-organization with applications in biology, physics, socio-economic sciences and engineering [4, 7, 25, 27, 34, 45, 48]. It describes a scenario with a group of $N \in \mathbb{N}$ agents, each of them having his/her own opinion or assessment of a certain quantity, represented by the real vector $x_i \in \mathbb{R}^d$, $i \in \{1, 2, \ldots, N\}$, with $d \geq 1$. Each of the agents communicates with all others and revises his/her own opinion based on a weighted average of the other agents' opinions. If this is done continuously in time, we have $x_i = x_i(t)$ for $t \geq 0$ and

$$\dot{x}_i(t) = \sum_{j=1}^{N} \psi_{ij}(t)(x_j(t) - x_i(t)), \qquad i \in [N], \tag{1}$$

where here and in the sequel we denote $[N] := \{1, 2, \ldots, N\}$. The communication weights $\psi_{ij} = \psi_{ij}(t)$ measure the intensity of the influence between agents depending on the dissimilarity of their opinions. This intensity, also called *bounded confidence*, is usually assumed to decay with increasing discord. In the classical setting [22] the communication weights are given by

$$\psi_{ij}(t) := \frac{1}{N-1}\psi(|x_j(t) - x_i(t)|), \tag{2}$$

*Computer, Electrical and Mathematical Sciences & Engineering, King Abdullah University of Science and Technology, 23955 Thuwal, KSA. jan.haskovec@kaust.edu.sa

with the nonnegative bounded *influence function* $\psi : [0, \infty) \to [0, \infty)$, also called *communication rate*. We adopt the assumption that $\psi(s) \leq 1$ for all $s \geq 0$, which implies

$$\sum_{j \neq i} \psi_{ij} \leq 1 \qquad \text{for all } i \in [N].$$

Let us note that this assumption is taken without loss of generality, since, as long as ψ is a globally bounded function, it can be achieved by an appropriate rescaling of time.

As pointed out in [32] for the second-order version of (1), aka the Cucker-Smale system [11, 12], the scaling by $1/N$ in (2) has the drawback that the dynamics of an agent is modified by the total number of agents even if its dynamics is only significantly influenced by a few nearby agents. Therefore, [32] proposed to normalize the communication weights relative to the influence of all other agents, without involving explicit dependence on their number,

$$\psi_{ij}(t) := \frac{\psi(|x_j(t) - x_i(t)|)}{\sum_{\ell \neq i} \psi(|x_\ell(t) - x_i(t)|)}. \tag{3}$$

The normalization removes the aforementioned drawback, but still allows for a meaningful passage to the mean field limit $N \to \infty$, see [32, 33]. It also induces the following property of the normalized weights (3),

$$\sum_{j \neq i} \psi_{ij} = 1 \qquad \text{for all } i = 1, \cdots, N, \tag{4}$$

so that the term on the right-hand side of (1) is a convex combination of the vectors $x_j(t) - x_i(t)$.

The main question is whether the dynamics (1) of continuously evolving opinions will tend to an (asymptotic) emergence of one or more *opinion clusters* formed by agents with (almost) identical opinions. *Global consensus* is the state where all agents have the same opinion, i.e., $x_i = x_j$ for all $i, j \in [N]$. Observe that the classical communication weights (2) possess the symmetry property $\psi_{ij} = \psi_{ji}$, which implies the conservation of the mean opinion $\frac{1}{N} \sum_{i=1}^{N} x_i$ along the dynamics induced by (1). Consequently, if the system reaches a global consensus, its value is determined by the mean opinion of the initial datum. On the other hand, the normalization in (3) destroys the symmetry of the communication weights and the global consensus value, if reached, cannot be easily inferred from the initial datum and can be seen as an emergent property of the system.

Various aspects of the consensus behavior of various modifications of (1) have been studied in, e.g., [2, 3, 5, 6, 30, 31, 32, 24, 46, 47]. For many applications in biological and socio-economical systems or control problems (for instance, swarm robotics [17, 43, 44]), it is natural to include a time delay in the model (1). From the modeling point of view, it is reasonable to consider two sources of delay:

- **transmission-type delay,** which reflects the fact that each agent may receive information from its surroundings with a certain time lag due to finite speed of information transmission and/or perception. Therefore, agent i with opinion $x_i(t)$ receives at time $t > 0$ the information about the opinion of agent j in the form $x_j(t - \tau)$. In general, the length of the delay τ may depend both on the state of the system and on the external conditions, i.e., it may be different for each pair (i, j), it may vary in time or follow a certain distribution. For simplicity of the exposition, we consider a globally constant delay $\tau > 0$ in this paper. Then, system (1) is replaced by

$$\dot{x}_i(t) = \sum_{j \neq i} \psi_{ij}(t)(x_j(t - \tau) - x_i(t)), \qquad i \in [N]. \tag{5}$$

  The communication rates $\psi_{ij} = \psi_{ij}(t)$ depend on the history of the trajectories in a suitable way, as discussed below.

- **Reaction-type delay,** which takes into account the time needed for the agents to react to the information that they already received and processed. In this case, the agent i's reaction executed at time t

is based upon the state of the system, including agent i's state, at time $t-\tau$. Again, the delay length τ may, in general, depend on both internal and external factors, however, we shall restrict to the simple case of a fixed, constant delay $\tau > 0$. Then, system (1) is replaced by

$$\dot{x}_i(t) = \sum_{j \neq i} \psi_{ij}(t)(x_j(t-\tau(t)) - x_i(t-\tau)), \qquad i \in [N], \tag{6}$$

where again the communication rates $\psi_{ij} = \psi_{ij}(t)$ depend on the history of the trajectories.

In both cases the system is equipped with the initial datum

$$x_i(t) = x_i^0(t), \qquad i = 1, \cdots, N, \quad t \in [-\tau, 0], \tag{7}$$

with prescribed continuous trajectories $x_i^0 \in \mathscr{C}([-\tau, 0])$, $i = 1, \cdots, N$.

The choice of a particular expression for the communication rates $\psi_{ij} = \psi_{ij}(t)$ in (5) and (6) should be made depending on modeling considerations. Again, there are two generic types of communication rates:

- **Classical communication rates,** inspired by (2), may take the form

$$\psi_{ij}(t) = \frac{1}{N-1}\psi(|x_j(t-\tau) - x_i(t)|) \tag{8}$$

for the system (5) with transmission-type delay, or

$$\psi_{ij}(t) = \frac{1}{N-1}\psi(|x_j(t-\tau) - x_i(t-\tau)|) \tag{9}$$

for the reaction-type delay system (6). However, we shall not set down any particular expression for ψ_{ij} in the sequel. Instead, we characterize the classical communication weights by the property

$$\sum_{j \neq i} \psi_{ij}(t) \leq 1 \qquad \text{for all } t \geq 0 \text{ and } i \in [N]. \tag{10}$$

Note that (10) is obviously verified by both the expressions above (recal the generic assumption $\psi \leq 1$, that can be made without loss of generality).

- **Normalized communication rates,** inspired by (3), may take the form

$$\psi_{ij}(t) := \frac{\psi(|x_j(t-\tau) - x_i(t)|)}{\sum_{\ell \neq i} \psi(|x_\ell(t-\tau) - x_i(t)|)} \tag{11}$$

for the system (5) with transmission-type delay, or

$$\psi_{ij}(t) := \frac{\psi(|x_j(t-\tau) - x_i(t-\tau)|)}{\sum_{\ell \neq i} \psi(|x_\ell(t-\tau) - x_i(t-\tau)|)} \tag{12}$$

for the reaction-type delay system (6). Again, we do not fix any particular form in the sequel. Instead, we characterize the normalized communication weights by the property

$$\sum_{j \neq i} \psi_{ij}(t) = 1 \qquad \text{for all } t \geq 0 \text{ and } i \in [N]. \tag{13}$$

Of course, (13) is a special case of (10). However, it turns the terms on the right-hand side of (5) and, resp., (6), into convex combinations of the vectors $x_j(t-\tau) - x_i(t)$ and, resp., $x_j(t-\tau) - x_i(t-\tau)$. This peculiar convexity property facilitates application of powerful analytical tools that provide stronger results and/or more elegant proofs, as we shall demonstrate below.

Regardless of the particular form of the communication rates we shall consider, we adopt the following *global positivity assumption* for the whole paper: If, for some $R > 0$ and $T > 0$, $|x_i(t) - x_j(t)| \leq R$ for all $t \in [-\tau, T]$ and all $i, j \in [N]$, then there exists $\underline{\psi} = \underline{\psi}(R) > 0$ such that

$$\psi_{ij}(t) \geq \frac{\underline{\psi}}{N-1} \quad \text{for all } t \in [-\tau, T] \text{ and all } i, j \in [N]. \tag{14}$$

Let us note that this assumption is satisfied by all the candidates (8)–(12) proposed above as soon as the influence function ψ is globally positive. Let us also note that the global positivity of ψ is obviously necessary for global consensus to be reached in general. Indeed, if $\psi(r) = 0$ for some $r > 0$, then one would have steady states formed by two clusters of particles at distance r apart.

In order to guarantee existence of global C^1-solutions of (5) and (6), we adopt the generic assumption that all $\psi_{ij} = \psi_{ij}(t)$ are continuous functions of time. This is verified by all the expressions (8)–(12) as soon as $\psi \in \mathscr{C}([0, \infty))$. Finally, let us stress that we do not impose any monotonicity property of ψ.

An important property that the consensus systems may or may not have is the *symmetry* and the resulting conservation of the mean

$$\overline{x}(t) := \frac{1}{N} \sum_{i=1}^{N} x_i(t). \tag{15}$$

Clearly, the transmission-type delay system (5) does not, in general, conserve the mean even if one chooses symmetric communication rates $\psi_{ij} = \psi_{ji}$. In contrast, the reaction-type system (6) conserves $\overline{x}(t)$ as soon as the communication rates are symmetric, like those given by (9). On the other hand, the normalized weights (12) destroy the symmetry. This is reflected in the methods that we shall use to prove asymptotic consensus for the respective systems. For the symmetric case we shall apply an ℓ^2-approach, working with quadratic fluctuations of x_i around the mean, while for the nonsymmetric case an ℓ^∞-approach is more suitable. The conservation of the mean (15) is a significant property that allows us to obtain stronger results for the symmetric case.

The aim of this paper is to provide a systematic overview of sufficient conditions for global asymptotic consensus in the systems (5) and (6) and the corresponding analytical tools. Some of the results are somehow scattered in the recent literature and we feel that a unifying perspective is beneficial. We also provide some new results and proofs. In particular, we shall provide the following four theorems:

- Theorem 1 for the system (5) with communication rates satisfying (10),
- Theorem 2 for (5) with normalized communication rates (13),
- Theorem 3 for the system (6) with symmetric communication rates,
- Theorem 4 for (6) with nonsymmetric communication rates (13).

The proofs are more or less closely following [19] and [18] for the system (5). The proof of Theorem 3 is an adaptation of the method developed in [14] for a second-order system with noise. The result for (6) with normalized (nonsymmetric) communication rates and its proof is based on an adaptation of the method developed in [20].

The paper is organized as follows. In Section 2 we consider the special case of (5) and (6) with two agents only, $N = 2$, and normalized communication weights. In this setting the systems reduce to single linear delay equations with known asymptotic behavior. They shall provide us intuition on what type of qualitative behavior we may expect from the solutions of the transmission-type delay consensus system (5) in contrast to the reaction-type system (6). In Section 3 we formulate the four Theorems about the asymptotic consensus for the systems (5) and (6) with the two types of communication rates discussed above. We also explain the main ideas and methods of the proofs. In Section 4 we give an overview of previous literature on the topic with brief comments on their results and methods. In Section 5 we prove an auxiliary Lemma of Gronwall-Halanay type that shall be of later use. Finally, in Sections 6 and 7 we provide proofs of the four theorems formulated in Section 3.

2 Toy models: $N = 2$

To gain an intuition on what type of qualitative behavior we may expect from the solutions of the delay consensus systems (5) and (6), we consider the special case with two agents only, $N = 2$, in the spatially one-dimensional setting $d = 1$. The goal of this section is to understand the different impact of the transmission-type versus the reaction-type delay on the consensus dynamics. Therefore, we consider agents with opinions $x_1(t)$, $x_2(t)$, and, to simplify the situation even further, we choose the normalized communication weights. Indeed, the convexity property (13) immediately implies $\psi_{12} = \psi_{21} = 1$.

With the above choices, the transmission-type delay system (5) reduces to the linear system

$$\begin{aligned}\dot{x}_1(t) &= x_2(t-\tau) - x_1(t),\\ \dot{x}_2(t) &= x_1(t-\tau) - x_2(t).\end{aligned}$$

Defining $w := x_1 - x_2$, we have

$$\dot{w}(t) = -w(t-\tau) - w(t). \tag{16}$$

Assuming a solution of the form $w(t) = e^{\xi t}$ with complex $\xi \in \mathbb{C}$, we obtain the characteristic equation

$$\xi + e^{-\xi\tau} + 1 = 0.$$

A simple inspection reveals that all roots ξ have negative real part, which implies that all solutions $w(t)$ of (16) tend to zero as $t \to \infty$. Therefore, we have the asymptotic consensus $\lim_{t\to\infty} x_1(t) - x_2(t) = 0$, for any value of the delay τ. This leads us to the intuitive expectation that all solutions of the general system with transmission-type delay (5) should converge to global consensus as $t \to \infty$, regardless of the length of the delay τ. Indeed, in Section 6 we shall prove this hypothesis for (5) with both the classical communication rates characterized by (10) and the normalized rates (13).

The situation is fundamentally different for the reaction-type delay in (6), which for $N = 2$ with (13) turns into

$$\begin{aligned}\dot{x}_1(t) &= x_2(t-\tau) - x_1(t-\tau),\\ \dot{x}_2(t) &= x_1(t-\tau) - x_2(t-\tau).\end{aligned}$$

In this case we obtain the delay negative feedback equation for $w := x_1 - x_2$,

$$\dot{w}(t) = -2w(t-\tau). \tag{17}$$

An analysis of the corresponding characteristic equation

$$\xi + 2\tau e^{-\xi} = 0,$$

with $\xi \in \mathbb{C}$, reveals that:

- If $0 < 2\tau < e^{-1}$, then $u = 0$ is asymptotically stable.
- If $e^{-1} < 2\tau < \pi/2$, then $u = 0$ is asymptotically stable, but every nontrivial solution of (17) is oscillatory.
- If $2\tau > \pi/2$, then $u = 0$ is unstable.

In fact, if $2\tau < e^{-1}$, the solutions subject to the constant initial datum never oscillate and tend to zero as $t \to \infty$. If 2τ becomes larger than e^{-1} but smaller than $\pi/2$, the nontrivial solutions must oscillate (i.e., change sign infinitely many times as $t \to \infty$), but the oscillations are damped and vanish as $t \to \infty$. Finally, for $2\tau > \pi/2$ the nontrivial solutions oscillate with unbounded amplitude as $t \to \infty$. We refer to Chapter 2 of [40] and [15] for more details.

The different types of qualitative behavior of solutions of (16) versus (17) can be intuitively understood by noting that the instantaneous negative feedback term $-w(t)$ in (16) has a stabilizing effect, while the delay feedback term $-w(t-\tau)$ induces oscillations, and this effect becomes stronger with longer delays. Thus, the stabilizing effect of the instantaneous term $-w(t)$ in the right-hand side of (16) is stronger than the destabilizing effect of the delay term $-w(t-\tau)$, regardless of the delay length τ. In (17) the stabilizing effect is not present, and therefore, for large enough delays, the solution develops oscillations with divergent amplitude as $t \to \infty$. An analogous intuition applies to the general Hegselmann-Krause systems with transmission (5) and reaction (6) type delay. In (5) the instantaneous negative feedback term $-x_i(t)$ has a stabilizing effect, which for any value of $\tau > 0$ is stronger than the destabilization induced by $x_j(t-\tau(t))$. On the other hand, in (6) the stabilization by $-x_i(t)$ does not take place, so that one can expect asymptotic consensus only for short enough delays $\tau > 0$.

Let us finally note that the traditional tool in study of asymptotic behavior of consensus-type models is the quadratic fluctuation, which for $N = 2$ reads simply $w(t)^2$. Then, for (16) we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}w(t)^2 = -w(t-\tau)w(t) - w(t)^2,$$

and for (17),

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}w(t)^2 = -2w(t-\tau)w(t).$$

From here we clearly see the inherent analytical difficulty imposed by the presence of delay: namely, that the term $w(t-\tau)w(t)$ does not have a definite sign as soon as oscillations occur. Consequently, the quadratic fluctuation does not need to be a monotonically decaying function of time. The goal of this paper is to present a systematic overview of methods that can be used to overcome this difficulty for the Hegselmann-Krause models (5) and (6).

3 Discrete description: Overview of results and methods

To formulate our main results, let us introduce the spatial diameter $d_x = d_x(t)$ of the agent group,

$$d_x(t) := \max_{i,j\in[N]} |x_i(t) - x_j(t)|. \tag{18}$$

The term *global asymptotic consensus* for the systems (5) and (6) is then defined as the property

$$\lim_{t\to\infty} d_x(t) = 0. \tag{19}$$

3.1 Transmission-type delay

As announced in Section 2, we expect global asymptotic consensus to be achieved for the system with transmission-type delay (5) for any $\tau > 0$, and for all initial data. Indeed, we have the following result:

Theorem 1. *Let the communication weights ψ_{ij} verify* (10) *and the global positivity property* (14)*. Then all solutions of* (5) *reach global asymptotic consensus in the sense of* (19)*.*

We shall present a proof of Theorem 1 in Section 6.1, based on the direct approach developed in [19]. It is an ℓ^∞-type proof working with the group diameter $d_x = d_x(t)$ defined in (18). We first consider the spatially one-dimensional setting and derive a uniform bound on the group diameter in terms of the initial datum. Then, we derive an explicit estimate on the diameter shrinkage on finite time intervals. By an iterative argument we then conclude asymptotic convergence of the group diameter to zero, i.e., consensus finding. The disadvantage of this method is that it does not provide convergence rates.

An improved result can be obtained assuming that the communication weights satisfy the convexity property (13), e.g., the normalized communication weights (11). In this case we have exponential convergence with calculable rate.

Theorem 2. *Let the influence function be continuous and strictly positive on* $[0, \infty)$ *and let the weights* ψ_{ij} *verify* (13)*. Then all solutions of* (5) *reach global asymptotic consensus as defined by* (19)*. The decay of* $d_x = d_x(t)$ *to zero is exponential with calculable rate.*

The proof of Theorem 2, developed in [18], will be presented in Section 6.2. It is again an ℓ^∞-type proof, and is significantly more elegant than the rather technical proof of Theorem 1 due to the benefit of the convexity property (13) of the communication weights. It consists of three ingredients. First, we prove nonexpansivity of the agent group, i.e., a uniform bound on the position radius of the agents. Second, we establish a convexity argument, fundamentally based on (13), which provides a simple delay differential inequality. Finally, we employ a Gronwall-Halanay-type inequality [16] to prove that its solutions decay to zero exponentially, with rate that can be calculated as a root of a nonlinear algebraic equation.

3.2 Reaction-type delay

Recalling the discussion in Section 2, we expect global asymptotic consensus to be achieved only for small enough values of $\tau > 0$ for the system (6) with reaction-type delay. Consequently, we can only hope to derive a sufficient condition, in terms of $\tau > 0$, that guarantees asymptotic consensus for all solutions. Moreover, due to the destabilizing effect of the reaction-type delay, as explained in Section 2, we need to adopt some additional/more restrictive assumptions than in the case of transmission-type delay.

Our first result assumes that the communication weights are symmetric, i.e., $\psi_{ij}(t) = \psi_{ji}(t)$ for all $i, j \in [N]$ and $t \geq 0$. Moreover, we impose the upper bound

$$\psi_{ij}(t) \leq \frac{1}{N-1}, \tag{20}$$

which is slightly more restrictive than (10). Note that both the symmetry condition and the upper bound (20) are verified by the classical weights (9), or, equally well by (8), if $\psi \leq 1$, which we assume without loss of generality.

Theorem 3. *Let the communication weights* ψ_{ij} *verify* (20)*, the global positivity* (14) *and the symmetry property* $\psi_{ij} = \psi_{ji}$*. If* $\tau \leq 1/2$*, then all solutions of* (6) *reach global asymptotic consensus as defined by* (19)*.*

The proof of Theorem 3 is an adaptation of the method developed in [14] and will be presented in Section 7.1. It is an ℓ^2-type proof, working with the quadratic fluctuation of the agent positions, and is based on a construction of an appropriate Lyapunov functional. It exploits the important consequence of the symmetry of the communication weigths, namely, the conservation of the mean $\overline{x}(t) := \frac{1}{N}\sum_{i=1}^N x_i(t)$ along the solutions of (6). Indeed, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{x}(t) = \frac{1}{N}\sum_{i=1}^N\sum_{j=1}^N \psi_{ij}(t)(x_j(t-\tau) - x_i(t-\tau)) = 0. \tag{21}$$

The drawback of the proof is that, similarly as the proof of Theorem 1, it does not provide exponential convergence to consensus.

Our second result works for non-symmetric communication weights satisfying (10). It provides exponential convergence to consensus with calculable rate, however, at the expense of a tighter restriction on the length of $\tau > 0$ and of the following assumption on the initial datum,

$$\max_{i\in[N]} |\dot{x}_i(t)| \leq d_x^0 \qquad \text{for almost all } t \in (-\tau, 0), \tag{22}$$

where we denoted

$$d_x^0 := \max_{t\in[-\tau,0]} d_x(t), \tag{23}$$

with $d_x = d_x(t)$ defined in (18). Let us note that (22) is trivially satisfied in the generic case of constant initial datum on $[-\tau, 0]$. Moreover, we need to tighten the global positivity assumption (14) in the sense that the lower bound on the communication weights ψ_{ij} depends continuously on the diameter of the group d_x. In particular, we assume that there exists a positive continuous function $\Psi \in \mathscr{C}([0,\infty))$ such that

$$\psi_{ij}(t) \geq \frac{\Psi(d_x(t-\tau))}{N-1} \quad \text{for all } t \geq 0 \text{ and all } i,j \in [N]. \tag{24}$$

This is satisfied by the classical weights (9) as long as $\psi > 0$, by choosing

$$\Psi(d_x) := \min_{s\in[0,d_x]} \psi(s). \tag{25}$$

For the normalized weights (12) we have

$$\psi_{ij}(t) = \frac{\psi(|x_j(t-\tau) - x_i(t-\tau)|)}{\sum_{\ell\neq i} \psi(|x_\ell(t-\tau) - x_i(t-\tau)|)} \geq \frac{\psi(|x_j(t-\tau) - x_i(t-\tau)|)}{N-1},$$

since, without loss of generality, $\psi \leq 1$. Consequently, (24) is again satisfied by choosing (25).

Theorem 4. *Let the communication rates ψ_{ij} satisfy* (10) *and* (24)*, and let the initial datum x^0 satisfy* (22)*. Denoting $\underline{\psi}^0 := (N-1)\min_{i,j\in[N]} \psi_{ij}(0)$, assume that*

$$4\tau < \underline{\psi}^0. \tag{26}$$

Then there exists a unique $C \in (0, \underline{\psi}^0)$ such that

$$\underline{\psi}^0 - C = 4e^{C\tau}\,\frac{e^{C\tau}-1}{C},$$

and the following estimate holds along solutions of (6)–(7)*,*

$$d_x(t) \leq d_x^0 e^{-Ct} \qquad \textit{for all } t \geq 0. \tag{27}$$

The proof of Theorem 4 is is an adaptation of the method developed in [20] for a second-order system, aka the Cucker-Smale model. It is based on a derivation of a delay integro-differential inequality and an application of an appropriately generalized Gronwall-Halanay-type inequality. We will present it in Section 7.2.

4 Overview of previous literature

Exponential convergence to global consensus as $t \to \infty$ for the system (5) with time varying delay $\tau = \tau(t)$ has been proved in [10], taking an ℓ^∞-approach and constructing a Lyapunov functional that is a sum of the group diameter d_x and an integral term. The result holds under a set of conditions requiring smallness of the maximal time delay in relation to the decay speed of the influence function and the fluctuation of the initial datum. A variant of the proof for (5) with distributed time delay was given in [36], again a smallness assumption on the maximal delay. These results are suboptimal since, as we state in Theorems 1 and 2, the transmission-type delay system (5) reaches asymptotic consensus without any restriction on τ, and for all initial data. Compared to Theorem 1, the advantage of [10, 36] is that it provides exponential convergence. Theorem 2 does establish exponential convergence too, however, only for the convex interaction weights (13).

In [37] the authors studied convergence to consensus of the Cucker-Smale model with time-varying delays, which can be considered a second-order version of the system (5). However, the authors imposed an assumption of a priori boundedness of the communication rates from below, which essentially reduces the model to first-order, i.e., a variant of (5). The main result of [37] is convergence to consensus under a smallness

assumption on the maximal time delay, which again is suboptimal in the context of transmission-type delay (cf. Theorem 1). The authors consider both symmetric interaction weights like (2), where they apply an ℓ^2-approach, and nonsymmetric weights like (3), where they take an ℓ^∞-approach. In both cases the proof is based on a construction of a suitable Lyapunov functional. In [38] the optimal flocking result has been finally extablished for the Cucker-Smale model with transmission-type delay, where no smallness assumption on the delay has to be imposed.

The convexity argument (see Lemma (5) below) that we shall apply for the case of normalized communication weights satisfying (13) is an adaptation of the notion of coefficient of ergodicity introduced by Dobrushin [13] in the context of Markov chains. It was later used to quantify the contractivity in first-order, discrete-time models of opinion dynamics (without delay) in [26]. In [23] the class of *scrambling matrices* was introduced, which can be characterized as stochastic matrices of communication weights with sufficient connectivity between agents. Then, the convexity argument, exploiting the stochasticity of the matrix, is used to provide a bound on the contractivity of the system, which ultimately leads to a proof of asymptotic consensus. A similar argument was applied to second-order models (without delay) in [33].

Second-order variants of (5) and (6) were studied in [28, 14, 21, 8, 9]. In particular, in [28] and [8] the Cucker-Smale model with transmission-type delay and normalized communication rates (11) was considered and asymptotic flocking was proved under a smallness condition on the delay, related to the decay properties of the influence function ψ and the velocity diameter of the initial datum. In [28] the authors study (5) with (11). They generalize the technique of estimating "maximal action of antisymmetric matrices", developed in [32], to the case with delay. However, this approach does not benefit from the convexity property (13) of the normalized communication weights and, therefore, leads to suboptimal results. The method of proof in [8] is more direct, based on deriving a system of dissipative delay differential inequalities and constructing a Lyapunov functional.

Models of consensus finding with delay have been extensively studied in the Engineering community, with results going back to [35]. We provide only a short and very incomplete compilation of references here. For linear problems, where the communication weights ψ_{ij} are fixed, stability criteria based on the frequency approach and on Lyapunov-Krasovskii techniques were derived in [39]. Proofs of delay-independent consensus and flocking in nonlinear networks with multiple time-varying delays under mild assumptions were provided in [41, 42]. These approaches are based on ℓ^∞-estimates obtained using fundamental concepts from the non-negative matrix theory and a Lyapunov-Krasovskii functional. Consensus over directed static networks with arbitrary finite communications delays was studied in [29].

5 A Gronwall-Halanay-type lemma

We provide a proof of a Gronwall-Halanay-type lemma, which establishes exponential decay of a nonnegative function, given that it satisfies a certain integro-differential inequality. The variant we present here is a slight generalization of [8, Lemma 2.5] and [18, Lemma 3.3].

Lemma 1. *Fix $\tau > 0$ and let $u \in \mathscr{C}([-2\tau, \infty))$ be a nonnegative continuous function with piecewise continuous derivative on $(0, \infty)$, such that for almost all $t > 0$ the differential inequality is satisfied,*

$$\frac{\mathrm{d}}{\mathrm{d}t} u(t) \leq \alpha \int_0^\tau u(t - s - \tau)\, \mathrm{dP}(s) - \beta u(t), \tag{28}$$

where P is a probability measure on $[0, \tau]$ and $0 < \alpha < \beta$. Then there exists $C \in (0, \beta - \alpha)$ such that

$$\beta - C = \alpha \int_0^\tau e^{C(s+\tau)}\, \mathrm{dP}(s), \tag{29}$$

and the estimate holds

$$u(t) \leq \left(\max_{s \in [-2\tau, 0]} u(s) \right) e^{-Ct} \qquad \textit{for all } t \geq 0. \tag{30}$$

Proof. It is easily seen that, for fixed $\tau > 0$, the right-hand side of (29) is an unbounded increasing function of $C > 0$, and that its limit as $C \to 0$ is α. On the other hand, the left-hand side of (29) is decreasing and equal to β for $C = 0$. Consequently, (29) is uniquely solvable whenever $\alpha < \beta$. Then we have

$$\beta - C = \alpha \int_0^\tau e^{C(s+\tau)} \, \mathrm{dP}(s) > \alpha,$$

so that $C \in (0, \beta - \alpha)$.

Let us denote

$$\bar{u} := \max_{s \in [-2\tau, 0]} u(s),$$

and define the functions $v = v(t)$ and $w = w(t)$ as

$$v(t) := \begin{cases} \bar{u} & \text{for } t \le 0, \\ u(t) & \text{for } t > 0, \end{cases} \qquad w(t) := \begin{cases} \bar{u} & \text{for } t \le 0, \\ \bar{u} e^{-Ct} & \text{for } t > 0. \end{cases}$$

Note that $v = v(t)$ verifies (28) for almost all $t > 0$. For any fixed $\lambda > 1$ set

$$\mathscr{S}_\lambda := \Big\{ t \ge 0 : v(s) \le \lambda w(s) \quad \text{for} \quad s \in [0, t] \Big\}.$$

Since $0 \in \mathscr{S}_\lambda$, $T_\lambda := \sup \mathscr{S}_\lambda \ge 0$ exists. We claim that

$$T_\lambda = \infty \qquad \text{for any } \lambda > 1.$$

For contradiction, assume $T_\lambda < \infty$ for some $\lambda > 1$. Due to the continuity of $v = v(t)$ for $t > 0$, there exists some $T_\lambda^* > T_\lambda$ such that v is differentiable at T_λ^* and

$$v(T_\lambda^*) > \lambda w(T_\lambda^*), \qquad \frac{\mathrm{d}}{\mathrm{d}t} v(T_\lambda^*) > \lambda \frac{\mathrm{d}}{\mathrm{d}t} w(T_\lambda^*). \tag{31}$$

Moreover, T_λ^* can be chosen such that $T_\lambda^* - \tau < T_\lambda$. Then we have, by construction, $v(s - \tau) \le \lambda w(s - \tau)$ for all $s \le T_\lambda^*$, and from (28) it follows

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} v(T_\lambda^*) &\le \alpha \int_0^\tau v(T_\lambda^* - s - \tau) \, \mathrm{dP}(s) - \beta v(T_\lambda^*) \\ &< \alpha\lambda \int_0^\tau w(T_\lambda^* - s - \tau) \, \mathrm{dP}(s) - \beta \lambda w(T_\lambda^*), \end{aligned}$$

where we also used (31). Next, noting that, trivially, $w(t) \le \bar{u} e^{-Ct}$ for all $t \in \mathbb{R}$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} v(T_\lambda^*) < \alpha\lambda\bar{u} \int_0^\tau e^{-C(T_\lambda^* - s - \tau)} \, \mathrm{dP}(s) - \beta\lambda w(T_\lambda^*),$$

and using the definition of $w(T_\lambda^*)$,

$$\frac{\mathrm{d}}{\mathrm{d}t} v(T_\lambda^*) < \left(\alpha \int_0^\tau e^{C(s+\tau)} \, \mathrm{dP}(s) - \beta \right) \lambda w(T_\lambda^*).$$

Finally, assumption (29) gives

$$\frac{\mathrm{d}}{\mathrm{d}t} v(T_\lambda^*) < -C\lambda w(T_\lambda^*) = \lambda \frac{\mathrm{d}}{\mathrm{d}t} w(T_\lambda^*),$$

which is a contradiction to the second part of (31).

We conclude that, for every $\lambda > 1$, $T_\lambda = \infty$ and $u(t) \le \lambda w(t)$ for all $t \ge 0$. Passing to the limit $\lambda \to 1$ yields the claim (30). □

We shall apply Lemma 1 twice in this paper: First in Section 6.2, where we choose the probability measure P to be a Dirac delta concentrated at zero. We then use it again in Section 7.2, this time with $\mathrm{P}(s) := \tau^{-1} \chi_{[0,\tau]}(s)$, where $\chi_{[0,\tau]}(s)$ denotes the characteristic function of the interval $[0, \tau]$.

6 Asymptotic consensus with transmission-type delay

A significant property of the system (5) with transmission-type delay is that one can derive a uniform in time bound on the radius $R_x = R_x(t)$ of the agent group,

$$R_x(t) := \max_{i\in[N]} |x_i(t)|, \tag{32}$$

in terms of the radius of the initial datum, defined as

$$R_x^0 := \max_{t\in[-\tau,0]} R_x(t).$$

In particular, we have the following result.

Lemma 2. *Let the communication weights satisfy* (10)*. Then, along the solutions of* (5)*, the diameter* $R_x = R_x(t)$ *defined in* (32) *satisfies*

$$R_x(t) \le R_x^0 \qquad \textit{for all } t \ge 0.$$

This result, combined with the global positivity property (14), directly implies that the communication weights $\psi_{ij} = \psi_{ij}(t)$ are uniformly bounded below by $\frac{\underline{\psi}}{N-1}$ with some $\underline{\psi} > 0$. We note that such a uniform a-priori bound is not available for the system (6) with reaction-type delay.

The proof of Lemma 2 (or its variant for variable delay) was given in [8, 10, 18]. Due to the fundamental role the result plays in the proof of asymptotic consensus, we provide it here for the sake of the reader.

Proof. Let us fix some $\varepsilon > 0$. We shall prove that for all $t \ge 0$

$$R_x(t) < R_x^0 + \varepsilon. \tag{33}$$

Obviously, $R_x(0) \le R_x^0$, so that by continuity, (33) holds on the maximal interval $[0, T)$ for some $T > 0$. For contradiction, let us assume that $T < +\infty$. Then we have

$$\lim_{t\to T-} R_x(t) = R_x^0 + \varepsilon. \tag{34}$$

However, for any $i = 1, \dots, N$, we have

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|x_i(t)|^2 &= \sum_{j\ne i} \psi_{ij}(t)\left[x_j(t-\tau) - x_i(t)\right]\cdot x_i(t) \\ &= \sum_{j\ne i} \psi_{ij}(t)\left[x_j(t-\tau)\cdot x_i(t) - |x_i(t)|^2\right]. \end{aligned}$$

By definition, we have $|x_j(t-\tau)| < R_x^0 + \varepsilon$ for all $j \in [N]$ and $t < T$, so that with an application of the Cauchy-Schwarz inequality we arrive at

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|x_i(t)|^2 &\le \sum_{j\ne i} \psi_{ij}(t)\left[\left(R_x^0 + \varepsilon\right)|x_i(t)| - |x_i(t)|^2\right] \\ &\le \left(R_x^0 + \varepsilon\right)|x_i(t)| - |x_i(t)|^2, \end{aligned}$$

where we used (10) in the second line. Now, if $|x_i(t)| \ne 0$, we use the identity $\frac{1}{2}\frac{\mathrm{d}|x_i(t)|^2}{\mathrm{d}t} = |x_i(t)|\frac{\mathrm{d}|x_i(t)|}{\mathrm{d}t}$ and divide the above inequality by $|x_i(t)|$. On the other hand, if $|x_i(t)| \equiv 0$ on an open subinterval of $(0, T)$, then $\frac{\mathrm{d}|x_i(t)|}{\mathrm{d}t} \equiv 0 \le R_x^0 + \varepsilon - |x_i(t)|$ on this subinterval. Thus, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}|x_i(t)| \le R_x^0 + \varepsilon - |x_i(t)| \quad \text{for almost all } t \in (0, T),$$

which implies

$$|x_i(t)| \leq \big(|x_i(0)| - (R_x^0 + \varepsilon)\big)\, e^{-t} + R_x^0 + \varepsilon \qquad \text{for } t < T.$$

Consequently, with $|x_i(0)| \leq R_x^0$,

$$\lim_{t \to T-} \max_{1 \leq i \leq N} |x_i(t)| \leq -\varepsilon e^{-T} + R_x^0 + \varepsilon < R_x^0 + \varepsilon,$$

which is a contradiction to (34). We conclude that, indeed, $T = \infty$, and complete the proof by taking the limit $\varepsilon \to 0$. □

6.1 Proof of Theorem 1

The proof of Theorem 1 was established in [19] and we provide its main ideas here. We first work in the spatially one-dimensional setting $d = 1$, i.e., $x_i(t) \in \mathbb{R}$. The 1D result is then easily generalized to the multi-dimensional situation by considering arbitrary one-dimensional projections of (5). We remind the reader that in Theorem 1 we only assumed the validity of the uniform bound (10) for the communication weights, together with the global positivity property (14).

In the one-dimensional setting, one can establish the following analogue of Lemma 2.

Lemma 3. *Let $d = 1$ and define*

$$m := \min_{i \in [N]} \min_{t \in [-\tau, 0]} x_i^0(t), \qquad M := \max_{i \in [N]} \max_{t \in [-\tau, 0]} x_i^0(t). \tag{35}$$

Let the communication weights satisfy (10)*. Then, along the solutions of* (5)*,*

$$m \leq x_i(t) \leq M$$

for all $i \in [N]$ and all $t \geq 0$.

Proof. The proof is similar to the proof of Lemma 2 and can be found in [19, Lemma 1]. □

Due to the global positivity assumption (14), Lemma 3 implies that there exists $\underline{\psi} > 0$ such that

$$\psi_{ij}(t) \geq \frac{\underline{\psi}}{N-1} \qquad \text{for all } t \geq 0 \text{ and } i, j \in [N]. \tag{36}$$

The following result establishes the fundament of the proof of asymptotic consensus. It shows that on finite time intervals the spatial diameter of the agent group shrinks by an explicit multiplicative factor.

Lemma 4. *Let the assumptions of Lemma 3 be verified and, moreover, assume $0 < m \leq M$. Then, along the solutions of* (5)*, we have for all $i \in [N]$ and $t \in [5\tau, 6\tau]$,*

$$m + \frac{\Gamma}{2}(M - m) \leq x_i(t) \leq M - \frac{\Gamma}{2}(M - m),$$

where

$$\Gamma := \left(1 - e^{-\frac{\underline{\psi}\tau}{N-1}}\right)^2 (1 - e^{-\sigma}) e^{-6\tau} \frac{\underline{\psi}}{N-1},$$

with $\underline{\psi}$ given by (36) *and*

$$\sigma := \min\left\{\tau, \frac{M - m}{2M}\right\}.$$

The proof is rather technical, based on direct estimation of the individual trajectories of the agents. We refer to [19] for its details and proceed directly with the proof of Theorem 1. Let us remind that we still consider the one-dimensional setting $d = 1$. We apply Lemma 4 iteratively on time intervals of total length 7τ. Let us for $k \in \mathbb{N}$ denote $\mathscr{I}_k := [(6k-1)\tau, 6k\tau]$ and

$$m_k := \min_{i\in[N]} \min_{t\in\mathscr{I}_k} x_i(t), \qquad M_k := \max_{i\in[N]} \max_{t\in\mathscr{I}_k} x_i(t).$$

Clearly, $m_0 = m$ and $M_0 = M$, with m and M given by (35). Moreover, let us introduce the notation $D_k := M_k - m_k$ and

$$\Gamma_k := \left(1 - e^{-\frac{\underline{\psi}\tau}{N-1}}\right)^2 (1 - e^{-\sigma_k}) e^{-6\tau} \frac{\underline{\psi}}{N-1},$$

with $\sigma_k := \min\left\{\tau, \frac{M_k - m_k}{2M_k}\right\}$. Note that $\Gamma_k \in (0,1)$ as long as $M_k > m_k$.

An application of Lemma 4 gives

$$m_0 + \frac{\Gamma_0}{2}(M_0 - m_0) \le x_i(t) \le M_0 - \frac{\Gamma_0}{2}(M_0 - m_0) \qquad \text{for } t \in \mathscr{I}_1.$$

We thus have

$$D_1 = M_1 - m_1 \le (1 - \Gamma_0)(M_0 - m_0) = (1 - \Gamma_0) D_0.$$

Iterating Lemma 4 gives

$$D_{k+1} \le (1 - \Gamma_k) D_k \qquad \text{for all } k \in \mathbb{N}.$$

Due to Lemma 3 we have $M_k \le M$ for all $k \in \mathbb{N}$, so that denoting $\overline{\sigma}(D) := \min\left\{\tau, \frac{D}{2M}\right\}$, we have $\sigma_k \ge \overline{\sigma}(D_k)$ for all $k \in \mathbb{N}$. Then, with

$$\overline{\Gamma}(D) := \left(1 - e^{-\frac{\underline{\psi}\tau}{N-1}}\right)^2 (1 - e^{-\overline{\sigma}(D)}) e^{-6\tau} \frac{\underline{\psi}}{N-1},$$

we have $\Gamma_k \ge \overline{\Gamma}(D_k)$ for all $k \in \mathbb{N}$ and

$$D_{k+1} \le \left(1 - \overline{\Gamma}(D_k)\right) D_k.$$

Clearly, the sequence $\{D_k\}_{k\in\mathbb{N}}$ is a nonnegative decreasing sequence, and denoting its limit D, the limit passage $k \to \infty$ in the above inequality gives $\overline{\Gamma}(D) \le 0$, which immediately implies $D = 0$. To conclude the proof of Theorem 1 in the one-dimensional setting, we apply the uniform bound of Lemma 3, i.e., if the agent group is contained in the spatial interval $[m_k, M_k]$ for $t \in \mathscr{I}_k$ then it remains contained in the same interval for all future times.

Finally, generalization of the proof to the spatially multi-dimensional setting is facilitated by the observation that the claims of both Lemma 3 and Lemma 4 can be trivially adapted to projections of the trajectories $x_i = x_i(t)$ to arbitrary one-dimensional subspaces of $\mathbb{R}^d$. Indeed, for an arbitrary fixed vector $\xi \in \mathbb{R}^d$, we replace (5) with the projected system

$$\frac{\mathrm{d}}{\mathrm{d}t}(x_i(t) \cdot \xi) = \sum_{j\ne i} \psi_{ij}(t)(x_j(t-\tau) - x_i(t)) \cdot \xi, \qquad i = 1, \dots, N,$$

and by a simple adaptation of the above proofs we obtain

$$\lim_{t\to\infty} (x_i(t) - x_j(t)) \cdot \xi = 0.$$

for all $i, j \in [N]$. We conclude by choosing ξ's as the basis vectors of $\mathbb{R}^d$.

6.2 Proof of Theorem 2

The proof goes along the lines of [18] and is fundamentally based on the convexity property (13) of the communication weights that is assumed in Theorem 2. In particular, we have the following result, whose proof can be found in [18, Lemma 3.2].

Lemma 5. *Let $N \geq 3$ and $\{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ be any set of vectors in $\mathbb{R}^d$ and denote d_x its diameter,*

$$d_x := \max_{1 \leq i,j \leq N} |x_i - x_j|.$$

Fix $i, k \in [N]$ such that $i \neq k$ and let $\eta_j^i \geq 0$ for all $j \in [N] \setminus \{i\}$, and $\eta_j^k \geq 0$ for all $j \in [N] \setminus \{k\}$, such that

$$\sum_{j \neq i} \eta_j^i = 1, \qquad \sum_{j \neq k} \eta_j^k = 1.$$

Let $\mu \geq 0$ be such that

$$\mu \geq \min \left\{ \min_{j \neq i} \eta_j^i, \ \min_{j \neq k} \eta_j^k \right\} \geq 0.$$

Then

$$\left| \sum_{j \neq i} \eta_j^i x_j - \sum_{j \neq k} \eta_j^k x_j \right| \leq (1 - (N-2)\mu) d_x.$$

In this section we shall apply Lemma 1 with the probability measure P being a Dirac delta concentrated at zero. This gives the following result.

Corollary 1. *Let $u \in \mathscr{C}([-\tau, \infty))$ be a nonnegative continuous function with piecewise continuous derivative on $(0, \infty)$, such that for some constants $0 < \alpha < \beta$ the differential inequality is satisfied,*

$$\frac{\mathrm{d}}{\mathrm{d}t} u(t) \leq \alpha u(t - \tau) - \beta u(t) \qquad \text{for almost all } t > 0.$$

Then there exists a unique solution $C \in (0, \beta - \alpha)$ of the equation

$$\beta - C = \alpha e^{C\tau} \tag{37}$$

and the estimate holds

$$u(t) \leq \left(\max_{s \in [-\tau, 0]} u(s) \right) e^{-Ct} \qquad \text{for all } t \geq 0.$$

Proof. Use Lemma 1 with $\mathrm{P}(s) := \delta(s)$. □

We are now prepared to prove Theorem 2.

Proof. Due to the continuity of the trajectories $x_i = x_i(t)$, there is an at most countable system of open, mutually disjoint intervals $\{\mathscr{I}_\sigma\}_{\sigma \in \mathbb{N}}$ such that

$$\bigcup_{\sigma \in \mathbb{N}} \overline{\mathscr{I}_\sigma} = [0, \infty)$$

and for each $\sigma \in \mathbb{N}$ there exist indices $i(\sigma)$, $k(\sigma)$ such that

$$d_x(t) = |x_{i(\sigma)}(t) - x_{k(\sigma)}(t)| \quad \text{for } t \in \mathscr{I}_\sigma.$$

Then, using the abbreviated notation $i := i(\sigma)$, $k := k(\sigma)$, we have for every $t \in \mathscr{I}_\sigma$,

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}d_x(t)^2 &= (\dot{x}_i(t) - \dot{x}_k(t)) \cdot (x_i(t) - x_k(t)) \\
&= \left(\sum_{j\neq i}\psi_{ij}(t)(x_j(t-\tau) - x_i(t)) - \sum_{j\neq k}\psi_{kj}(t)(x_j(t-\tau) - x_k(t))\right) \cdot (x_i(t) - x_k(t)) \\
&= \left(\sum_{j\neq i}\psi_{ij}(t)x_j(t-\tau) - \sum_{j\neq k}\psi_{kj}(t)x_j(t-\tau)\right) \cdot (x_i(t) - x_k(t)) - |x_i(t) - x_k(t)|^2,
\end{aligned}$$

where we used the convexity property of the normalized weights (4). Lemma 2 together with the global positivity assumption (14) provides some $\underline{\psi} \in (0,1)$ such that

$$\psi_{ij}(t) \geq \frac{\underline{\psi}}{N-1} \qquad \text{for all } t \geq 0 \text{ and } i,j \in [N].$$

We then apply Lemma 5 with $\mu := \frac{\underline{\psi}}{N-1}$, which gives

$$\begin{aligned}
\left|\sum_{j\neq i}\psi_{ij}x_j(t-\tau) - \sum_{j\neq k}\psi_{kj}x_j(t-\tau)\right| &\leq (1-(N-2)\mu)d_x(t-\tau) \\
&= \left(1 - \frac{N-2}{N-1}\underline{\psi}\right) d_x(t-\tau)\,.
\end{aligned}$$

Consequently, with the Cauchy-Schwartz inequality we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}d_x(t)^2 \leq \left(1 - \frac{N-2}{N-1}\underline{\psi}\right) d_x(t-\tau)d_x(t) - d_x(t)^2,$$

which implies that for almost all $t > 0$,

$$\frac{\mathrm{d}}{\mathrm{d}t}d_x(t) \leq \left(1 - \frac{N-2}{N-1}\underline{\psi}\right) d_x(t-\tau) - d_x(t).$$

An application of Corollary 1 with $\alpha := 1 - \frac{N-2}{N-1}\underline{\psi}$ and $\beta := 1$ gives then the exponential decay

$$d_x(t) \leq \left(\max_{s\in[-\tau,0]} d_x(s)\right) e^{-Ct} \qquad \text{for } t \geq 0,$$

where C is the unique solution of (37), i.e.,

$$1 - C = \left(1 - \frac{N-2}{N-1}\underline{\psi}\right) e^{C\tau}.$$

□

7 Asymptotic consensus with reaction-type delay

In this section we provide proofs of Theorems 3 and 4 that give sufficient conditions for global asymptotic consensus in the system with reaction-type delay (6). Let us recall that, according to the discussion of the simplified model in Section 2, we cannot expect asymptotic consensus to take place for arbitrary delay lengths. Indeed, the sufficient conditions of Theorems 3 and 4 are formulated as smallness conditions on $\tau > 0$. This is related to the fact that, in contrast to the system with transmission-type delay (Lemma 2), no uniform bound on the radius of the agent group is available.

For the sake of legibility, in this section we shall use the shorthand notation $\widetilde{x}_j := x_j(t-\tau)$, while x_j means $x_j(t)$.

7.1 Proof of Theorem 3

Let us recall that Theorem 3 assumes symmetric interaction weights

$$\psi_{ij}(t) = \psi_{ji}(t) \qquad \text{for all } i, j \in [N] \text{ and } t \geq 0, \tag{38}$$

which implies the conservation of the mean (21) along the solutions of (6). Consequently, without loss of generality we may assume that

$$\overline{x}(t) = \frac{1}{N} \sum_{i=1}^{N} x_i(t) \equiv 0$$

for all $t \geq 0$. We then introduce the quadratic fluctuation of the positions around the mean,

$$X(t) := \frac{1}{2(N-1)} \sum_{i=1}^{N} |x_i|^2. \tag{39}$$

Moreover, for $t \geq 0$ we define the quantity

$$D(t) := \frac{1}{2(N-1)} \sum_{i=1}^{N} \sum_{j \neq i} \psi_{ij} |\widetilde{x}_j - \widetilde{x}_i|^2. \tag{40}$$

We first derive an estimate on the dissipation of the quadratic fluctuation $X = X(t)$ in terms of $D = D(t)$.

Lemma 6. *For any $\delta > 0$ we have, along the solutions of* (6) *with symmetric interaction weights* (38)*,*

$$\frac{\mathrm{d}}{\mathrm{d}t} X(t) \leq (\delta - 1) D(t) + \tau \delta^{-1} \int_{t-\tau}^{t} D(s) \, \mathrm{d}s \qquad \textit{for all } t > \tau. \tag{41}$$

Proof. With (6) we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} X(t) &= \frac{1}{N-1} \sum_{i=1}^{N} x_i \cdot \dot{x}_i \\
&= \frac{1}{N-1} \sum_{i=1}^{N} \sum_{j \neq i} \psi_{ij}(t) (\widetilde{x}_j - \widetilde{x}_i) \cdot x_i \\
&= \frac{1}{N-1} \sum_{i=1}^{N} \sum_{j \neq i} \psi_{ij}(t) (\widetilde{x}_j - \widetilde{x}_i) \cdot \widetilde{x}_i + \frac{1}{N-1} \sum_{i=1}^{N} \sum_{j \neq i} \psi_{ij}(t) (\widetilde{x}_j - \widetilde{x}_i) \cdot (x_i - \widetilde{x}_i).
\end{aligned}$$

For the first term of the right-hand side we apply the standard symmetrization trick (exchange of summation indices $i \leftrightarrow j$, noting the symmetry of $\psi_{ij} = \psi_{ji}$),

$$\frac{1}{N-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \psi_{ij}(t) (\widetilde{x}_j - \widetilde{x}_i) \cdot \widetilde{x}_i = -\frac{1}{2(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} \psi_{ij}(t) |\widetilde{x}_j - \widetilde{x}_i|^2 .$$

Therefore, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} X(t) = -\frac{1}{2(N-1)} \sum_{i=1}^{N} \sum_{j \neq i} \psi_{ij}(t) |\widetilde{x}_j - \widetilde{x}_i|^2 + \frac{1}{N-1} \sum_{i=1}^{N} \sum_{j \neq i} \psi_{ij}(t) (\widetilde{x}_j - \widetilde{x}_i) \cdot (x_i - \widetilde{x}_i).$$

For the last term we use the Young inequality with $\delta > 0$,

$$\begin{aligned}\left|\frac{1}{N-1}\sum_{i=1}^{N}\sum_{j\neq i}\psi_{ij}(t)(\widetilde{x}_j-\widetilde{x}_i)\cdot(x_i-\widetilde{x}_i)\right| &\leq \frac{\delta}{2(N-1)}\sum_{i=1}^{N}\sum_{j\neq i}\psi_{ij}(t)|\widetilde{x}_j-\widetilde{x}_i|^2+\frac{\delta^{-1}}{2(N-1)}\sum_{i=1}^{N}\sum_{j\neq i}\psi_{ij}(t)|x_i-\widetilde{x}_i|^2\\ &\leq \frac{\delta}{2(N-1)}\sum_{i=1}^{N}\sum_{j\neq i}\psi_{ij}(t)|\widetilde{x}_j-\widetilde{x}_i|^2+\frac{\delta^{-1}}{2(N-1)}\sum_{i=1}^{N}|x_i-\widetilde{x}_i|^2,\end{aligned}$$

where we used (10) for the second inequality. Hence,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}X(t) &\leq \frac{\delta-1}{2(N-1)}\sum_{i=1}^{N}\sum_{j\neq i}\psi_{ij}(t)|\widetilde{x}_j-\widetilde{x}_i|^2+\frac{\delta^{-1}}{2(N-1)}\sum_{i=1}^{N}|x_i-\widetilde{x}_i|^2\\ &= (\delta-1)D(t)+\frac{\delta^{-1}}{2(N-1)}\sum_{i=1}^{N}|x_i-\widetilde{x}_i|^2. \qquad (42)\end{aligned}$$

Next, for $t > \tau$ we use (6) to evaluate the difference $x_i - \widetilde{x}_i$,

$$x_i-\widetilde{x}_i=\int_{t-\tau}^{t}\dot{x}_i(s)\,\mathrm{d}s=\sum_{j\neq i}\int_{t-\tau}^{t}\psi_{ij}(s)(x_j(s-\tau)-x_i(s-\tau))\,\mathrm{d}s.$$

Taking the square and summing over $i \in [N]$ we arrive at

$$\begin{aligned}\sum_{i=1}^{N}|x_i-\widetilde{x}_i|^2 &= \sum_{i=1}^{N}\left|\sum_{j\neq i}\int_{t-\tau}^{t}\psi_{ij}(s)(x_j(s-\tau)-x_i(s-\tau))\,\mathrm{d}s\right|^2\\ &\leq (N-1)\sum_{i=1}^{N}\sum_{j\neq i}\left|\int_{t-\tau}^{t}\psi_{ij}(s)(x_j(s-\tau)-x_i(s-\tau))\,\mathrm{d}s\right|^2\\ &\leq \tau\sum_{i=1}^{N}\sum_{j\neq i}\int_{t-\tau}^{t}\psi_{ij}(s)|x_j(s-\tau)-x_i(s-\tau)|^2\,\mathrm{d}s\\ &= 2(N-1)\tau\int_{t-\tau}^{t}D(s)\,\mathrm{d}s.\end{aligned}$$

The first inequality is the Cauchy-Schwartz inequality for the summation term, i.e. $\left|\sum_{j\neq i}a_i\right|^2 \leq (N-1)\sum_{j\neq i}|a_i|^2$, and the second Cauchy-Schwartz inequality for the integral term, together with the bound $\psi_{ij} \leq 1/(N-1)$ provided by assumption (20). Inserting into (42) finally gives (41). □

We now provide the proof of Theorem 3.

Proof. For some $\lambda > 0$ we define the functional

$$\mathscr{L}(t) := X(t)+\lambda\int_{t-\tau}^{t}\int_{\theta}^{t}D(s)\,\mathrm{d}s\,\mathrm{d}\theta,$$

where $X = X(t)$ is the quadratic velocity fluctuation (39) and $D = D(t)$ defined in (40).

The time derivative of the second term in $\mathscr{L}(t)$ yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\lambda\int_{t-\tau}^{t}\int_{\theta}^{t}D(s)\,\mathrm{d}s\,\mathrm{d}\theta\right)=\lambda\tau D(t)-\lambda\int_{t-\tau}^{t}D(s)\,\mathrm{d}s.$$

Combining this with (41), with the choice $\lambda := \tau\delta^{-1}$ so that we eliminate the integral term, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{L}(t) \le \left(\delta - 1 + \tau^2\delta^{-1}\right) D(t) \qquad \text{for all } t \ge \tau.$$

Optimizing in $\delta > 0$ leads to $\delta := \tau$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathscr{L}(t) \le (2\tau - 1)\, D(t) \qquad \text{for all } t \ge \tau. \tag{43}$$

Therefore, we have

$$X(t) \le \mathscr{L}(t) \le \mathscr{L}(\tau) \qquad \text{for all } t \ge \tau,$$

and, recalling the conservation of the mean $\bar{x}(t)$, the uniform boundedness of $X(t)$ implies uniform boundedness of $|x_i - x_j|$ for all $i, j \in [N]$. Then, the global positivity assumption (14) gives some $\underline{\psi} > 0$ such that

$$\psi_{ij}(t) \ge \frac{\underline{\psi}}{N-1} \qquad \text{for all } t \ge \tau.$$

Then, with $\bar{x}(t) \equiv 0$, we have

$$D(t) = \frac{1}{2(N-1)} \sum_{i=1}^{N} \sum_{j \ne i} \psi_{ij} |\widetilde{x}_j - \widetilde{x}_i|^2 \ge \frac{\underline{\psi}}{2(N-1)^2} \sum_{i=1}^{N} \sum_{j \ne i} |\widetilde{x}_j - \widetilde{x}_i|^2 = 2\underline{\psi} X(t-\tau).$$

Integrating (43) on $[\tau, t]$ gives

$$\begin{aligned} X(t) \le \mathscr{L}(t) &\le \mathscr{L}(\tau) + (2\tau - 1) \int_\tau^t D(s)\,\mathrm{d}s \\ &\le \mathscr{L}(\tau) + 2\,(2\tau - 1)\,\underline{\psi} \int_0^{t-\tau} X(s)\,\mathrm{d}s. \end{aligned}$$

Taking limes superior as $t \to \infty$ in both sides, we have

$$2\,(1 - 2\tau)\,\underline{\psi} \int_0^\infty X(s)\,\mathrm{d}s + \limsup_{t\to\infty} X(t) \le \mathscr{L}(\tau), \tag{44}$$

so that $X = X(t)$ is integrable at infinity if $\tau < 1/2$. Barbalat's lemma [1] then states that $X(t) \to 0$ as $t \to \infty$ as soon as $\frac{\mathrm{d}}{\mathrm{d}t}X(t)$ is uniformly bounded, say, for $t > \tau$. This follows easily by an application of Lemma 6 with $\delta := 1$,

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) \le \tau \int_{t-\tau}^t D(s)\,\mathrm{d}s \le \tau^2 \max_{s\in[t-\tau,t]} D(s),$$

combined with (40) and (20),

$$\begin{aligned} D(t) &= \frac{1}{2(N-1)} \sum_{i=1}^{N} \sum_{j \ne i} \psi_{ij} |\widetilde{x}_j - \widetilde{x}_i|^2 \\ &\le \frac{1}{(N-1)^2} \sum_{i=1}^{N} \sum_{j \ne i} |\widetilde{x}_j|^2 + |\widetilde{x}_i|^2 \\ &\le 4X(t-\tau) \end{aligned}$$

and the observation that (44) implies uniform boundedness of $X = X(t)$. □

7.2 Proof of Theorem 4

Here we consider the system with reaction-type delay (6) with non-symmetric communication weights. The proof of asymptotic consensus presented in this section is new. We assume that the communication weights satisfy the global upper bound (10) and depend continuously on the diameter of the group in the sense of (25). We also impose the technical assumption (22) on the initial datum. For the sake of more elegant expressions, let us extend the definition of $d_x = d_x(t)$ to $t < 0$ as follows,

$$d_x(t) := \begin{cases} \max_{i,j\in[N]} |x_i(t) - x_j(t)| & \text{for } t > 0, \\ d_x^0 & \text{for } t \le 0, \end{cases} \tag{45}$$

where d_x^0 is given by (23).

Lemma 7. *Let the communication weights satisfy* (10) *and assume that there exists* $\underline{\psi} > 0$ *such that, for some* $T > 0$,

$$\psi_{ij}(t) \ge \frac{\underline{\psi}}{N-1} \qquad \text{for all } t \in [0,T] \text{ and } i,j \in [N]. \tag{46}$$

Moreover, let the initial datum verify (22). *Then, along the solutions of* (6),

$$\frac{\mathrm{d}}{\mathrm{d}t} d_x(t) \le 4 \int_{t-\tau}^{t} d_x(s-\tau)\mathrm{d}s - \underline{\psi} d_x(t)$$

for almost all $t \in (0,T)$.

Proof. Due to the continuity of the trajectories $x_i = x_i(t)$, there is an at most countable system of open, mutually disjoint intervals $\{\mathscr{I}_\sigma\}_{\sigma\in\mathbb{N}}$ such that

$$\bigcup_{\sigma\in\mathbb{N}} \overline{\mathscr{I}_\sigma} = [0,\infty)$$

and for each $\sigma \in \mathbb{N}$ there exist indices $i(\sigma)$, $k(\sigma)$ such that

$$d_x(t) = |x_{i(\sigma)}(t) - x_{k(\sigma)}(t)| \quad \text{for } t \in \mathscr{I}_\sigma.$$

Then, using the abbreviated notation $i := i(\sigma)$, $k := k(\sigma)$, we have for every $t \in \mathscr{I}_\sigma$,

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} d_x(t)^2 &= (\dot{x}_i - \dot{x}_k)\cdot(x_i - x_k) \\ &= \left(\sum_{j\ne i} \psi_{ij}(t)(\widetilde{x}_j - \widetilde{x}_i) - \sum_{j\ne k} \psi_{kj}(t)(\widetilde{x}_j - \widetilde{x}_k) \right) \cdot (x_i - x_k). \end{aligned} \tag{47}$$

We process the first term of the right-hand side as follows,

$$\begin{aligned} \sum_{j\ne i} \psi_{ij}(t)(\widetilde{x}_j - \widetilde{x}_i)\cdot(x_i - x_k) &= \sum_{j\ne i} \psi_{ij}(t)(\widetilde{x}_j - x_j + x_i - \widetilde{x}_i)\cdot(x_i - x_k) \\ &\quad + \sum_{j\ne i} \psi_{ij}(t)(x_j - x_i)\cdot(x_i - x_k). \end{aligned} \tag{48}$$

We estimate the difference $|\widetilde{x}_j - x_j|$ by

$$\begin{aligned}
|\widetilde{x}_j - x_j| &\leq \int_{t-\tau}^{t} |\dot{x}_j(s)|\,\mathrm{d}s \\
&\leq \int_{-[t-\tau]^-}^{0} |\dot{x}_j(s)|\,\mathrm{d}s + \int_{[t-\tau]^+}^{t} \sum_{\ell\neq j} \psi_{j\ell}(s)|x_\ell(s-\tau) - x_j(s-\tau)|\,\mathrm{d}s \\
&\leq \int_{-[t-\tau]^-}^{0} d_x^0\,\mathrm{d}s + \int_{[t-\tau]^+}^{t} \sum_{\ell\neq j} \psi_{j\ell}(s)\, d_x(s-\tau)\,\mathrm{d}s \\
&\leq \int_{t-\tau}^{t} d_x(s-\tau)\,\mathrm{d}s,
\end{aligned}$$

where for the third inequality we used the assumption (22) and for the last inequality we used the property (10), together with (45). Performing an analogous estimate for the term $|\widetilde{x}_i - x_i|$ and using the Cauchy-Schwartz inequality and, again, (10), we arrive at

$$\begin{aligned}
\sum_{j\neq i} \psi_{ij}(t)(\widetilde{x}_j - x_j + x_i - \widetilde{x}_i)\cdot(x_i - x_k) &\leq \sum_{j\neq i} \psi_{ij}(t)\left(|\widetilde{x}_j - x_j| + |\widetilde{x}_i - x_i|\right)|x_i - x_k| \\
&\leq 2\, d_x(t) \int_{t-\tau}^{t} d_x(s-\tau)\,\mathrm{d}s.
\end{aligned}$$

To estimate the second term of the right-hand side of (48), observe that, using the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
(x_j - x_i)\cdot(x_i - x_k) &= (x_j - x_k)\cdot(x_i - x_k) - |x_i - x_k|^2 \\
&\leq |x_i - x_k|\big(|x_j - x_k| - |x_i - x_k|\big) \leq 0,
\end{aligned}$$

since, by definition, $|x_j - x_k| \leq d_x = |x_i - x_k|$. Then, by assumption (46), we have

$$\sum_{j\neq i} \psi_{ij}(t)(x_j - x_i)\cdot(x_i - x_k) \leq \frac{\underline{\psi}}{N-1} \sum_{j=1}^{N} (x_j - x_i)\cdot(x_i - x_k).$$

Repeating the same steps for the second term of the right-hand side of (47), we finally arrive at

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} d_x(t)^2 &\leq 4\, d_x(t) \int_{t-\tau}^{t} d_x(s-\tau)\,\mathrm{d}s + \frac{\underline{\psi}}{N-1}\left(\sum_{j=1}^{N}(x_j - x_i)\cdot(x_i - x_k) - \sum_{j=1}^{N}(x_j - x_k)\cdot(x_i - x_k)\right) \\
&= 4\, d_x(t) \int_{t-\tau}^{t} d_x(s-\tau)\,\mathrm{d}s - \frac{N\underline{\psi}}{N-1}|x_i - x_k|^2 \\
&\leq 4\, d_x(t) \int_{t-\tau}^{t} d_x(s-\tau)\,\mathrm{d}s - \underline{\psi}\, d_x(t)^2.
\end{aligned}$$

Consequently, for almost all $t \in (0, T)$,

$$\frac{\mathrm{d}}{\mathrm{d}t} d_x(t) \leq 4 \int_{t-\tau}^{t} d_x(s-\tau)\,\mathrm{d}s - \underline{\psi}\, d_x(t).$$

□

In the sequel we shall apply the Gronwall-Halanay Lemma 1 with the probability measure $\mathrm{P}(s) := \tau^{-1}\chi_{[0,\tau]}(s)$, where $\chi_{[0,\tau]}(s)$ is the characteristic function of the interval $[0, \tau]$. This gives the following result.

Corollary 2. *Let $u \in \mathscr{C}([-2\tau, \infty))$ be a nonnegative continuous function with piecewise continuous derivative on $(0, \infty)$, such that for almost all $t > 0$ the differential inequality is satisfied,*

$$\frac{\mathrm{d}}{\mathrm{d}t} u(t) \leq \frac{\alpha}{\tau} \int_{t-\tau}^{t} u(s-\tau)\,\mathrm{d}s - \beta u(t).$$

Moreover, let $0 < \alpha < \beta$. Then there exists $C \in (0, \beta - \alpha)$ such that

$$\beta - C = \alpha e^{C\tau} \frac{e^{C\tau} - 1}{C\tau}, \tag{49}$$

and the estimate holds

$$u(t) \leq \left(\max_{s \in [-2\tau, 0]} u(s) \right) e^{-Ct} \qquad \textit{for all } t \geq 0.$$

Proof. Apply Lemma 1 with $\mathrm{P}(s) := \tau^{-1} \chi_{[0,\tau]}(s)$, noting that

$$\int_0^\tau e^{C(s+\tau)}\,\mathrm{dP}(s) = \frac{1}{\tau} \int_0^\tau e^{C(s+\tau)} \mathrm{d}s = e^{C\tau} \frac{e^{C\tau} - 1}{C\tau}.$$

□

We are now ready to carry out the proof of Theorem 4.

Proof. We define the set

$$\mathscr{S} := \left\{ T \geq 0;\ d_x(t) \leq d_x^0 \text{ for all } t \in [0, T] \right\}.$$

Obviously, $0 \in S$, so that there exists $T^* := \sup \mathscr{S}$. We claim that $T^* = \infty$. For contradiction, assume that $T^* < \infty$, then we have $d_x(t) \leq d_x^0$ for all $t \in [0, T^*]$. Since, by assumption (26),

$$4\tau < \underline{\psi}^0 = (N-1) \min_{i,j \in [N]} \psi_{ij}(0),$$

and the weights $\psi_{ij} = \psi_{ij}(t)$ depend continuously on the group diameter in the sense of (24), which in turn depends continuously on time along the solutions of (6), there exists $\varepsilon > 0$ and $\underline{\psi} > 4\tau$ such that

$$\underline{\psi} \leq (N-1) \min_{i,j \in [N]} \psi_{ij}(t) \qquad \text{for } t \in [0, T^* + \varepsilon].$$

Lemma 7 gives then

$$\frac{\mathrm{d}}{\mathrm{d}t} d_x(t) \leq 4 \int_{t-\tau}^{t} d_x(s-\tau)\mathrm{d}s - \underline{\psi} d_x(t),$$

and Corollary 2 with $\alpha := 4\tau$ and $\beta := \underline{\psi}$ yields

$$d_x(t) \leq d_x^0 e^{-Ct} \leq d_x^0 \qquad \text{for } t \in [0, T^* + \varepsilon]$$

with $C = C(\underline{\psi}) > 0$ given by (49). This is a contradiction to the choice of T^* being the supremum of $\mathscr{S}$.

Therefore, we have $d_x(t) \leq d_x^0$ for all $t \geq 0$ and, consequently, by (24),

$$\min_{i,j \in [N]} \psi_{ij}(t) \geq \frac{\underline{\psi}^0}{N-1} \qquad \text{for } t \geq 0.$$

An application of Lemma 7 and Corollary 2 with $\alpha := 4\tau$ and $\beta := \underline{\psi}^0$ finally gives (27). □

References

[1] I. Barbalat: *Systèmes d'équations différentielles d'oscillations nonlinéaires.* Rev. Math. Pures Appl. 4, 2 (1959), 267–270.

[2] A. Bhattacharyya, M. Braverman, B. Chazelle, and H.L. Nguyen: *On the convergence of the Hegselmann-Krause system.* Proceedings of the 4th conference on Innovations in Theoretical Computer Science, pp. 61–66. ACM, New York (2013).

[3] V.D. Blondel, J.M. Hendrickx and J.N. Tsitsiklis: *On Krause's multi-agent consensus model with state-dependent connectivity.* IEEE Trans. Autom. Control 54 (2009), 2586–2597.

[4] S. Camazine, J. L. Deneubourg, N.R. Franks, J. Sneyd, G. Theraulaz and E. Bonabeau: *Self-Organization in Biological Systems.* Princeton University Press, Princeton, NJ, 2001.

[5] C. Canuto, F. Fagnani and P. Tilli: *An Eulerian approach to the analysis of Krause's consensus models.* SIAM J. Control Optim. 50 (2012), 243–265.

[6] A. Carro, R. Toral, M. San Miguel: *The role of noise and initial conditions in the asymptotic solution of a bounded confidence, continuous-opinion model.* J. Stat. Phys. 151 (2013), 131–149.

[7] C. Castellano, S. Fortunato and V. Loreto: *Statistical physics of social dynamics.* Rev. Mod. Phys., 81, (2009), 591–646.

[8] Y.-P. Choi and J. Haskovec: *Cucker-Smale model with normalized communication weights and time delay.* Kinetic and Related Models 10 (2017), 1011-1033.

[9] Y.-P. Choi and J. Haskovec: *Hydrodynamic Cucker-Smale model with normalized communication weights and time delay..* SIAM J. Math. Anal., Vol. 51, No. 3 (2019), 2660–2685.

[10] Y.-P. Choi, A. Paolucci and C. Pignotti: *Consensus of the Hegselmann-Krause opinion formation model with time delay.* Math Meth Appl Sci. (2020), 1– 20.

[11] F. Cucker and S. Smale, Emergent behaviour in flocks, *IEEE T. on Automat. Contr.*, **52** (2007), 852–862.

[12] F. Cucker and S. Smale, On the mathematics of emergence, *Jap. J. Math.*, **2** (2007), 197–227.

[13] R. L. Dobrushin, *Central limit theorem for non-stationary Markov chains. I*, Theory Probab. Appl., 1:1 (1956), 65–80.

[14] R. Erban, J. Haskovec and Y. Sun, A Cucker-Smale model with noise and delay, *SIAM J. Appl. Math.*, **76** (2016), 1535–1557.

[15] I. Gyori and G. Ladas, *Oscillation Theory of Delay Differential Equations with Applications*, Oxford Science Publications, Clarendon Press, Oxford, 1991.

[16] A. Halanay, *Differential Equations: Stability, Oscillations, Time Lags.* Academic Press, New York London, 1966.

[17] H. Hamman: *Swarm Robotics: A Formal Approach.* Springer, 2018.

[18] J. Haskovec, *A simple proof of asymptotic consensus in the Hegselmann-Krause and Cucker-Smale models with renormalization and delay.* SIAM J. on Applied Dynamical Systems, 20:1 (2021), pp. 130–148.

[19] J. Haskovec, *Direct proof of unconditional asymptotic consensus in the Hegselmann-Krause model with transmission-type delay.* Bull. London Math. Soc. 53(5):1312-1323 (2021).

[20] J. Haskovec, *Flocking in the Cucker-Smale model with self-delay and nonsymmetric interaction weights.* J. Math. Anal. Appl. 514 (2022), pp. 126261.

[21] J. Haskovec and I. Markou, *Asymptotic flocking in the Cucker-Smale model with reaction-type delays in the non-oscillatory regime.* Kinetic and Related Models 13 (2020), 795–813.

[22] R. Hegselmann and U. Krause, *Opinion dynamics and bounded confidence models, analysis, and simulation.* J. Artif. Soc. Soc. Simul., 5, (2002), 1–24.

[23] R. Hegselmann and U. Krause: *Consensus and Fragmentation of Opinions with a Focus on Bounded Confidence.* The American Mathematical Monthly 126:8 (2019), 700–716.

[24] P.E. Jabin and S. Motsch: *Clustering and asymptotic behavior in opinion formation.* J. Differential Equations 257 (2014), 4165–4187.

[25] A. Jadbabaie, J. Lin and A. S. Morse: *Coordination of groups of mobile autonomous agents using nearest neighbor rules.* IEEE Trans. Automat. Control, 48, (2003), 988–1001.

[26] U. Krause: *A discrete nonlinear and non-autonomous model of consensus formation.* Communications in difference equations, 227-236 (2000).

[27] P. Krugman: *The Self Organizing Economy.* Blackwell Publishers, 1995.

[28] Y. Liu, and J. Wu, Flocking and asymptotic velocity of the Cucker-Smale model with processing delay, *J. Math. Anal. Appl.*, **415** (2014), 53–61.

[29] J. Lu, D. W. C. Ho and J. Kurths: *Consensus over directed static networks with arbitrary finite communications delays.* Phys. Rev. E, 80 (2009), 066121.

[30] S. Mohajer and B. Touri: *On convergence rate of scalar Hegselmann-Krause dynamics.* Proceedings of the IEEE American Control Conference (ACC) (2013).

[31] L. Moreau: *Stability of multiagent systems with time-dependent communication links.* IEEE Trans. Autom. Control 50(2), 169–182 (2005).

[32] S. Motsch and E. Tadmor: *A New Model for Self-organized Dynamics and Its Flocking Behavior.* J. Stat. Phys. 144 (2011).

[33] S. Motsch and E. Tadmor: *Heterophilious Dynamics Enhances Consensus.* SIAM review 56 (4), 577–621 (2014).

[34] G. Naldi, L. Pareschi and G. Toscani (eds.): *Mathematical Modeling of Collective behaviour in Socio-Economic and Life Sciences*, Series: Modelling and Simulation in Science and Technology, Birkhäuser, 2010.

[35] S.-I. Niculescu: *Delay Effects on Stability. A Robust Control Approach.* Springer-Verlag London, 2001.

[36] A. Paolucci: *Convergence to consensus for a Hegselmann-Krause-type model with distributed time delay.* Preprint (2020), `arxiv.org/abs/2005.08030`.

[37] C. Pignotti and E. Trelat: *Convergence to consensus of the general finite-dimensional Cucker-Smale model with time-varying delays.* Comm. Math. Sci. 16 (2018), 2053–2076.

[38] M. Rodriguez Cartabia, *Cucker-Smale model with time delay.* Discrete Contin. Dynam. Systems, 42, (2022), pp. 2409–2432.

[39] A. Seuret, V. Dimos, V. Dimarogonas and K.H. Johansson: *Consensus under Communication Delays.* Proceedings of the 47th IEEE Conference on Decision and Control, Cancun, Mexico, Dec. 9-11, 2008.

[40] H. Smith: *An Introduction to Delay Differential Equations with Applications to the Life Sciences.* Springer New York Dordrecht Heidelberg London, 2011.

[41] C. Somarakis and J. Baras: *A Simple Proof of the Continuous Time Linear Consensus Problem with Applications in Non-Linear Flocking Networks.* 2015 European Control Conference (ECC), 1546–1553.

[42] C. Somarakis and J. Baras: *Delay-independent convergence for linear consensus networks with applications to non-linear flocking systems.* In Proceedings of the 12th IFAC Workshop on Time Delay Systems, pp. 159–164, Ann Arbor (2015).

[43] K. Szwaykowska, I.B. Schwartz, L.M. Romero, C.R. Heckman, D. Mox and M. Ani Hsieh: *Collective motion patterns of swarms with delay coupling: theory and experiment.* Phys. Rev. E 93, 032307.

[44] G. Valentini: *Achieving Consensus in Robot Swarms: Design and Analysis of Strategies for the best-of-n Problem.* Springer, Studies in Computational Intelligence, Vol. 706, 2017.

[45] T. Vicsek and A. Zafeiris: *Collective motion.* Phys. Rep., 517 (2012), 71–140.

[46] C. Wang, Q. Li, W. E, B. Chazelle: *Noisy Hegselmann-Krause Systems: Phase Transition and the 2R-Conjecture.* J Stat Phys (2017) 166:1209–1225.

[47] E. Wedin and P. Hegarty: *A quadratic lower bound for the convergence rate in the one-dimensional Hegselmann-Krause bounded confidence dynamics.* Discret. Comput. Geom. 53(2), 478–486 (2015).

[48] H. Xu, H.Wang and Z. Xuan: *Opinion dynamics: a multidisciplinary review and perspective on future research.* Int. J. Knowl. Syst. Sci., 2, (2011), 72–91.