

Asymptotic behavior of the linear consensus model with delay and anticipation

Jan Haskovec*

Abstract

We study asymptotic behavior of solutions of the first-order linear consensus model with delay and anticipation, which is a system of neutral delay differential equations. We consider both the transmission-type and reaction-type delay that are motivated by modeling inputs. Studying the simplified case of two agents, we show that, depending on the parameter regime, anticipation may have both a stabilizing and destabilizing effect on the solutions. In particular, we demonstrate numerically that moderate level of anticipation generically promotes convergence towards consensus, while too high level disturbs it. Motivated by this observation, we derive sufficient conditions for asymptotic consensus in the multiple-agent systems, which are explicit in the parameter values delay length and anticipation level, and independent of the number of agents. The proofs are based on construction of suitable Lyapunov-type functionals.

Keywords: Asymptotic consensus, long-time behavior, delay, anticipation, neutral functional differential equations.

2010 MR Subject Classification: 34K05, 82C22, 34D05, 92D50.

1 Introduction

In this paper we study asymptotic behavior of the linear consensus model consisting of a group of $N \in \mathbb{N}$ agents, each of them having their own opinion or assessment of a certain quantity, represented by the real vector $x_i \in \mathbb{R}^d$, $i \in [N]$, with $d \geq 1$ the space dimension. Here and in the sequel we denote the set $[N] := \{1, 2, \dots, N\}$. Each of the agents communicates with all others and revises its own opinion based on a weighted average of the perceived other agents' opinions. If this is done continuously in time, we have $x_i = x_i(t)$ for $t \geq 0$ and

$$\dot{x}_i(t) = \sum_{j=1}^N \psi_{ij}(t)(\xi_j(t) - \xi_i(t)). \quad (1)$$

Here $\xi_j = \xi_j(t)$ represents the information that at time $t > 0$ is available about agent j 's opinion to all other agents. The communication weights $\psi_{ij} = \psi_{ij}(t)$ measure the intensity of the influence between agent i and agent j . The choice $\xi_j(t) := x_j(t)$ for all $j \in [N]$ and $\psi_{ij}(t) := \psi(|x_i(t) - x_j(t)|)$ gives the very well studied Hegselmann-Krause system [14], which is generic for modeling self-organized consensus in systems of interacting agents [2, 3, 16, 18, 22, 33, 34].

For many applications in biological and socio-economical systems or control problems (for instance, swarm robotics [10, 31, 32]), it is natural to include a time delay (lag) in the model reflecting the time needed for each agent to receive information from other agents, and/or to react to it. In this paper we also assume that the agents are able to anticipate the action of their conspecifics by measuring and extrapolating the

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

momentary rate of change of their state vectors. However, this information is as well subject to the time lag. 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 be random, following a certain distribution. For simplicity, in this paper we consider a fixed, globally constant delay $\tau > 0$. From the modeling point of view, it is reasonable to consider the following two types of systems.

- **System without self-delay (transmission-type delay)**, which reflects the situation when each agent receives 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 $x_j(t - \tau)$. Also, it is able to measure the rate of change of agent j 's opinion, which it receives as $\dot{x}_j(t - \tau)$, and extrapolates. Consequently, we have

$$\xi_j(t) := x_j(t - \tau) + \lambda\tau\dot{x}_j(t - \tau), \quad \xi_i(t) := x_i(t),$$

and (1) turns into

$$\dot{x}_i(t) = \sum_{j \neq i} \psi_{ij}(t)(x_j(t - \tau) + \lambda\tau\dot{x}_j(t - \tau) - x_i(t)), \quad i \in [N]. \quad (2)$$

The parameter $\lambda \geq 0$ denotes the strength of anticipation. The value $\lambda = 1$ corresponds to the Taylor expansion of the smooth curve $x_j = x_j(t)$ at time $t - \tau$ to first order; $\lambda > 1$ corresponds to (over)extrapolation, while $\lambda < 1$ can be seen as interpolation between $x_j(t - \tau)$ and $x_j(t - \tau) + \tau\dot{x}_j(t - \tau)$. Note that the summation in (2) runs through all $j \in [N]$ such that $j \neq i$. This prevents x_i from interacting with its own opinion from the past, which obviously is not justified from the modeling point of view.

- **System with self-delay (reaction-type delay)**, which takes into account the time needed for the agents to react to the information that they already received. In this case, the agent i 's reaction executed at time t is based upon the state of the system, including agent i 's own state, and its rate of change, at time $t - \tau$. We thus have

$$\xi_j(t) := x_j(t - \tau) + \lambda\tau\dot{x}_j(t - \tau), \quad \xi_i(t) := x_i(t - \tau) + \lambda\tau\dot{x}_i(t - \tau),$$

and (1) takes the form

$$\dot{x}_i(t) = \sum_{j=1}^N \psi_{ij}(t)(x_j(t - \tau) + \lambda\tau\dot{x}_j(t - \tau) - x_i(t - \tau) - \lambda\tau\dot{x}_i(t - \tau)), \quad i \in [N]. \quad (3)$$

In both cases the system is equipped with the initial datum

$$x_i(t) = x_i^0(t), \quad i \in [N], \quad t \in [-\tau, 0], \quad (4)$$

with prescribed continuously differentiable trajectories $x_i^0 \in C^1([-\tau, 0])$, $i \in [N]$. We note that both systems (2) and (3) fall into the class of neutral delay differential equations, see, e.g. [9, Chapter 9].

As already noted, in the classical Hegselmann-Krause bounded confidence model the communication weights $\psi_{ij} = \psi_{ij}(t)$ depend nonlinearly on the dissimilarity $|x_i(t) - x_j(t)|$ of the opinions of the respective agents. In this paper we resort to prescribing constant communication weights ψ_{ij} for all $i, j \in [N]$. This restriction, not unusual in previous literature [26, 28, 29], is motivated by the analytical approach that we shall employ. It is based on constructing Lyapunov functionals involving, for the system (2), terms of the type

$$x_i - \lambda\tau \sum_{j \neq i} \psi_{ij} x_j(t - \tau).$$

The time derivative of this term along the solutions of (2) reads

$$\frac{d}{dt} \left(x_i - \lambda \tau \sum_{j \neq i} \psi_{ij} x_j(t - \tau) \right) = \sum_{j \neq i} \psi_{ij} (x_j(t - \tau) - x_i(t)),$$

and the key observation is that it consists of derivative-free terms only, which can be suitably estimated. Similarly, for (3) we shall use

$$\frac{d}{dt} \left(x_i - \lambda \tau \sum_{j \neq i} \psi_{ij} (x_j(t - \tau) - x_i(t - \tau)) \right) = \sum_{j \neq i} \psi_{ij} (x_j(t - \tau) - x_i(t - \tau)).$$

However, the systems (2) and (3) with constant communication weights can be seen as linearizations of the respective nonlinear systems about their steady states. The sufficient conditions for the global asymptotic consensus in the linear systems that we shall present in Section 2 below can then be seen as conditions for local asymptotic stability of the respective nonlinear steady states. Further structural assumptions on the properties of the matrix $(\psi_{ij})_{i,j \in [N]}$ shall be specified in Section 2.

The main question in the context of consensus formation models of the type (1) is whether the dynamics of continuously evolving opinions tends to an (asymptotic) emergence of one or more opinion clusters [15] formed by agents with (almost) identical opinions. In case of global communications (i.e., $\psi_{i,j} > 0$ for all $i, j \in [N]$), one is interested in *global consensus*, which is the state where all agents have the same opinion $x_i = x_j$ for all $i, j \in [N]$. For the Hegselmann-Krause model with delay and its various modifications, this question has been studied extensively in numerous works, see, e.g., [4, 11, 12, 13, 19, 20, 23, 24, 25, 26, 28, 29]. Anticipation in the context of collective dynamics was introduced in [27], however, without delay or time lag. The authors studied large-time behavior of systems driven by radial potentials, which react to anticipated positions. As a special case, they considered the Cucker-Smale model of alignment [5, 6], proving the decisive role of anticipation in driving such systems with attractive potentials into velocity alignment and spatial concentration. They also studied the concentration effect near equilibrium for anticipation-based dynamics of pair of agents governed by attractive–repulsive potentials.

However, to our best knowledge, none of the previous works considered the first-order consensus model of the type (1) with delay and anticipation, i.e., no results are found in the literature concerning the asymptotic behavior of the systems (2) and (3) or their variants. The main goal of this paper is to fill this gap. It is known that the presence of delay generically causes destabilization of solutions through appearance of damped or undamped oscillations [9, 30, 8]. One may then conjecture that anticipation could oppose this effect by re-establishing stability in certain parameter regimes, which in the context of the systems (2), (3) means promotion of convergence of solutions towards global consensus. To gain insights into this hypothesis, we first study the simplified case of two-agent systems, where (2) and (3) reduce to single neutral delay differential equations. Combining known analytical results and numerical simulations, we shall demonstrate that moderate levels of anticipation indeed promote stability. In particular, for the system (2) with transmission-type delay, moderate anticipation reduces the amplitude of the oscillations of the solutions. For the system (3) with reaction-type delay we shall numerically identify a parameter range where anticipation promotes stability in otherwise unstable solutions through a Hopf-type bifurcation. Noting that anticipation may well have a destabilizing effect on the dynamics motivates the search for sufficient conditions that guarantee convergence towards global consensus in the systems (2) and (3). We shall provide such conditions explicitly in the parameter values $\lambda \geq 0$ and $\tau \geq 0$ in the respective systems with constant communication weights. These results, formulated as Theorems 1 and 2 below, are, to our best knowledge, new. The proofs will be based on construction of suitable Lyapunov functionals and, despite the linearity of the considered systems, highly nontrivial due to the presence of anticipation and delay terms.

This paper is organized as follows. In Section 2 we present our main results on the sufficient conditions for global asymptotic consensus for the systems (2) and (3), discussing in detail the assumptions on the communication weights that we adopt. In Section 3 we consider the simplified setting with two agents only, $N = 2$, where (2) and (3) reduce to single neutral delay differential equations. We then provide insights

into the dynamics of both systems by discussing known analytical results and presenting results of numerical simulations. In Section 4 we provide the proof of our main result on global asymptotic consensus for the transmission-type delay system (2), and in Section 5 we provide the proof for the reaction-type delay system (2).

2 Assumptions and main results

In this section we introduce the particular assumptions on the communication weights ψ_{ij} appearing in the systems (2) and (3). As explained in the Introduction, we shall work with fixed, time-independent weights in this paper, which facilitates analysis of the asymptotic behavior of the solutions using Lyapunov-type functionals. As the mathematical properties of the systems (2) and (3) differ, we shall introduce two different sets of assumptions on ψ_{ij} below. Then, we shall formulate two theorems providing sufficient conditions for the asymptotic stability of the trivial steady state, i.e., sufficient conditions for reaching asymptotic consensus.

Let us remark that both (2) and (3) are instances of linear neutral functional differential equations. Consequently, by standard results, see, e.g., [9, Chapter 9], they possess unique global solutions subject to the continuously differentiable initial datum (4).

For later use let us define the mean opinion

$$X(t) := \frac{1}{N} \sum_{i=1}^N x_i(t) \quad (5)$$

and the maximum opinion discord

$$d_x(t) := \max_{i,j \in [N]} |x_i(t) - x_j(t)|. \quad (6)$$

Clearly, global consensus corresponds to $d_x = 0$.

2.1 Transmission-type delay

For the system (2) without self-delay, we assume that the matrix $(\psi_{ij})_{i,j \in [N]}$ of communication weights is row-stochastic in the sense

$$\sum_{j \neq i} \psi_{ij} = 1 \quad \text{for all } i \in [N]. \quad (7)$$

As the diagonal elements ψ_{ii} do not appear in (2), we may set them to zero and treat the whole matrix $(\psi_{ij})_{i,j \in [N]}$ as row stochastic. Moreover, we assume that all non-diagonal elements are strictly positive,

$$\psi_{ij} > 0 \quad \text{for all } i \neq j. \quad (8)$$

Let us observe that the system (2), in general, does not conserve the mean (5), and even an eventual symmetry of the matrix $(\psi_{ij})_{i,j \in [N]}$, which we however do *not* assume in the sequel, would not guarantee conservation of the mean. Consequently, 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. As we shall see in Section 4, this also brings some additional difficulties for the analysis of the asymptotic stability of the global consensus state.

Furthermore, we assume that there exists some $T_0 > 0$ and fixed indices $I, J \in [N]$ such that

$$d_x(t) = |x_I(t) - x_J(t)| \quad \text{for } t \geq T_0. \quad (9)$$

Admittedly, the validity of this assumption is difficult to examine a priori, although our extensive numerical investigations suggest its generic plausibility. The necessity for adopting this assumption stems from the form of the Lyapunov functional that we shall construct in order to study the asymptotic behavior of system (2) and will be explained in detail in Remark 1 of Section 4.

Theorem 1. *Let $N \geq 3$ and let the matrix $(\psi_{ij})_{i,j \in [N]}$ of communication weights satisfy the row stochasticity (7) and positivity (8). Let the assumption (9) be satisfied and let*

$$\lambda\tau \leq 1. \tag{10}$$

Then all solutions of (2) subject to the initial datum (4) converge to a consensus state, i.e., $\lim_{t \rightarrow \infty} d_x(t) = 0$.

Theorem 1 gives sufficient conditions for reaching global consensus as $t \rightarrow \infty$ in the system (2) with at least three agents. As will be discussed in Section 3, for $N = 2$ asymptotic consensus is reached if and only if $|\lambda\tau| < 1$.

It is well known [13, Theorem 2.1] that without anticipation ($\lambda = 0$), system (2) with row-stochastic interaction matrix reaches asymptotic consensus unconditionally, i.e., for all initial data and for all values of the delay $\tau \geq 0$. Consequently, anticipation may only provide the benefit of reducing the amplitude of eventual oscillations of the solutions. This is indeed the case for small enough values of $\lambda > 0$, as we shall demonstrate numerically in Section 3.1. However, choosing λ too large leads to loss of stability. The sufficient condition (10) of Theorem 1 can therefore be understood as an explicit quantitative guarantee in terms of the parameter values λ and τ that stability is preserved.

2.2 Reaction-type delay

The system (3) with self-delay has the remarkable property that, if the matrix of communication weights $(\psi_{ij})_{i,j \in [N]}$ is symmetric, then (3) does conserve the mean (5). Then, if a global consensus is reached, it is determined by the initial condition through $X(0)$.

Moreover, we assume that the matrix $(\psi_{ij})_{i,j \in [N]}$ is irreducible. Note that a matrix is irreducible if and only if it represents a (weighted) connectivity matrix of a strongly connected graph (directed graph is called strongly connected if there is a path in each direction between each pair of its vertices). Clearly, this is a minimal necessary condition for reaching global consensus.

Theorem 2. *Let the matrix $(\psi_{ij})_{i,j \in [N]}$ of communication weights be symmetric and irreducible. Let*

$$(1 + \lambda)\tau < 1/2. \tag{11}$$

Then all solutions of (3) subject to the initial datum (4) converge to the consensus state $X(0)$, i.e.,

$$\lim_{t \rightarrow \infty} x_i(t) = X(0) \quad \text{for all } i \in [N],$$

with the mean $X(0)$ defined in (5).

Condition (11) of Theorem 2 can be read as a smallness condition on $\lambda > 0$ relative to a given $\tau < 1/2$. As we shall observe numerically in Section 3.2 for $N = 2$, there is a parameter regime where solutions with moderate anticipation, i.e., with $\lambda > 0$, are stable (converge to consensus), while solutions without anticipation, i.e., $\lambda = 0$, are unstable. Obviously, Theorem 2 does not capture this phenomenon. The reason is that its proof is based on a construction of a Lyapunov functional, which naturally leads to a sufficient condition that imposes smallness of λ and τ . In fact, we are not aware of any analytical results that would explain this stabilization-by-anticipation phenomenon. On the other hand, too large values of $\lambda > 0$ have the opposite effect, i.e., they cause loss of stability of solutions that are stable without anticipation. Therefore, the value of the statement of Theorem 2 lies in guaranteeing preservation of stability.

3 The simplified case - two agents

To gain insight into the type of dynamics produced by the systems (2) and (3), we consider their maximally simplified versions with two agents only, $N = 2$. Note that then (7) dictates $\psi_{12} = \psi_{21} = 1$ for the transmission-type delay system (2), and we shall use the same values also for the reaction-type delay system (3).

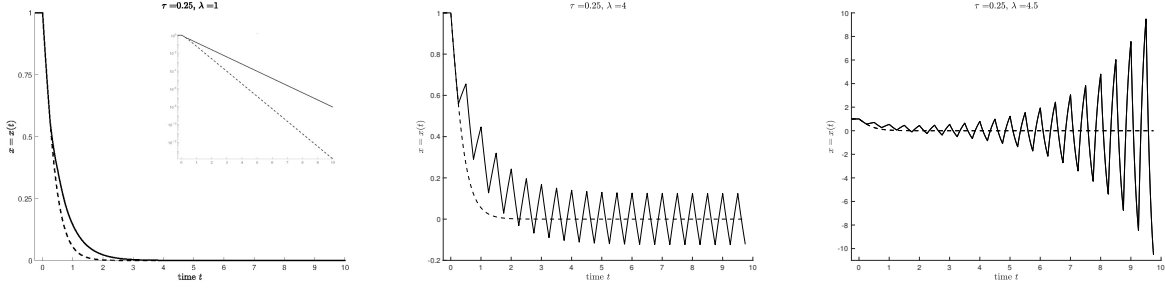


Figure 1: Numerical integration of (12) subject to the constant initial datum $x^0 \equiv 1$, with $\tau = 0.25$ and $\lambda = 1$ (solid line in left panel), $\lambda = 4$ (solid line in middle panel) and $\lambda = 4.5$ (solid line in right panel). For comparison, numerical solution of (12) with $\tau = 0.25$ and $\lambda = 0$ (no anticipation) is indicated by dashed line in all three panels. The inserted plot in the left panel is in logarithmic scale (vertical axis).

3.1 Transmission delay with $N = 2$

The system (2) with $N = 2$ and $\psi_{12} = \psi_{21} = 1$ reads

$$\begin{aligned}\dot{x}_1 &= \tilde{x}_2 + \lambda\tau\dot{\tilde{x}}_2 - x_1, \\ \dot{x}_2 &= \tilde{x}_1 + \lambda\tau\dot{\tilde{x}}_1 - x_2.\end{aligned}$$

Denoting $x := x_1 - x_2$, we have

$$\dot{x} = -\tilde{x} - \lambda\tau\dot{\tilde{x}} - x, \quad (12)$$

equipped with the initial datum $x(t) = x^0(t)$ for $t \in [-\tau, 0]$, with $x^0 \in C^1([-\tau, 0])$. Clearly, consensus in the context of (12) corresponds to $x = 0$. It is known that the zero steady state of (12) is globally stable if and only if $|\lambda\tau| < 1$, see [9, Chapter 9.9]. In other words, for a given $\tau > 0$, anticipation in (12) leads to consensus, for all initial data $x^0 \in C^1([-\tau, 0])$, if and only if $\lambda < \tau^{-1}$. Note that without anticipation (i.e., $\lambda = 0$), the solution always converges to the consensus $x = 0$, for any length of the delay $\tau \geq 0$.

We illustrate this fact by a numerical simulation. We discretize (12) on an equidistant mesh with time step $\Delta t > 0$, chosen such that τ is an integer multiple of Δt . Implicit Euler discretization of \dot{x} and $\dot{\tilde{x}}$ leads then to a difference equation that can be numerically integrated. We show results of numerical integration with the constant initial datum $x^0 \equiv 1$. In Fig. 1 we consider the case $\tau := 0.25$, where the solution of (12) without anticipation (i.e., $\lambda := 0$) decreases monotonically without oscillations. With $\lambda := 1$, which corresponds to Taylor expansion of the respective term, the solutions still decays monotonically, but the decay is slower, as is evident from the inserted logarithmic plot. Choosing $\lambda := 4$, which gives the borderline case $\lambda\tau = 1$, oscillations appear that convergence to a periodic solution. Finally, after crossing the threshold $\lambda\tau = 1$ by choosing $\lambda := 4.5$, diverging oscillations appear. In Fig. 2 we choose $\tau := 1.25$, where the solution of (12) without anticipation develops oscillations (but still converges to zero as $t \rightarrow \infty$). We first choose $\lambda := 0.2$ to demonstrate that small enough anticipation may actually decrease the amplitude of the oscillations. However, already the choice $\lambda := 0.6$ leads to an increase in the amplitude compared with the no-anticipation case. Finally, the Taylor anticipation $\lambda := 1$ crosses the boundary of stability $\lambda\tau = 1$ and produces a diverging solution.

3.2 Reaction delay with $N = 2$

The system (3) with $N = 2$ and $\psi_{12} = \psi_{21} = 1$ reads

$$\begin{aligned}\dot{x}_1 &= \tilde{x}_2 + \lambda\tau\dot{\tilde{x}}_2 - (\tilde{x}_1 + \lambda\tau\dot{\tilde{x}}_1), \\ \dot{x}_2 &= \tilde{x}_1 + \lambda\tau\dot{\tilde{x}}_1 - (\tilde{x}_2 + \lambda\tau\dot{\tilde{x}}_2).\end{aligned}$$

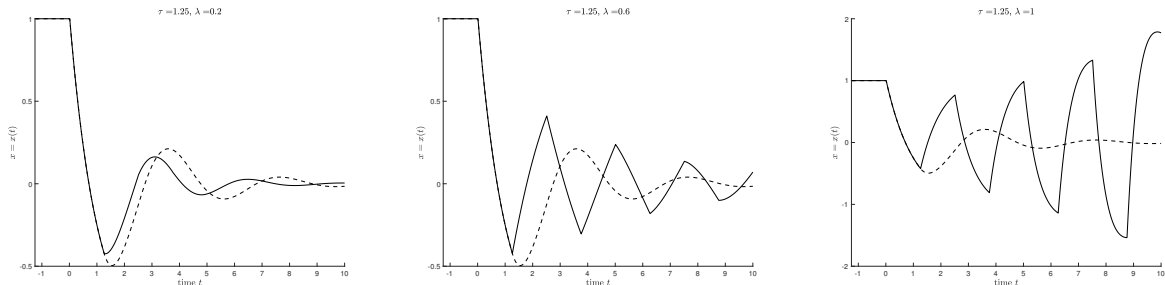


Figure 2: Numerical integration of (12) subject to the constant initial datum $x^0 \equiv 1$, with $\tau = 1.25$ and $\lambda = 0.2$ (solid line in left panel), $\lambda = 0.6$ (solid line in middle panel) and $\lambda = 1$ (solid line in right panel). For comparison, numerical solution of (12) with $\tau = 1.25$ and $\lambda = 0$ (no anticipation) is indicated by dashed line in all three panels.

Denoting $x := x_1 - x_2$, we have

$$\dot{x} = -2\tilde{x} - 2\lambda\tau\dot{\tilde{x}}, \quad (13)$$

subject to the initial datum $x(t) = x^0(t)$ for $t \in [-\tau, 0]$, with $x^0 \in C^1([-\tau, 0])$. Although stability of equations of the type (13) has been studied in the literature, see, e.g., [17], no explicit analytical condition for stability is known to us that would cover the case $\lambda\tau > 0$ in (13). If there is no anticipation, i.e., $\lambda := 0$, then (13) reduces to

$$\dot{x}(t) = -2x(t - \tau), \quad (14)$$

and an analysis of the corresponding characteristic equation reveals that:

- If $0 < 2\tau < e^{-1}$, then $x = 0$ is asymptotically stable.
- If $e^{-1} < 2\tau < \pi/2$, then $x = 0$ is asymptotically stable, but every nontrivial solution of (14) is oscillatory (i.e., changes sign infinitely many times as $t \rightarrow \infty$).
- If $2\tau > \pi/2$, then $x = 0$ is unstable and nontrivial solutions oscillate with unbounded amplitude as $t \rightarrow \infty$.

We refer to Chapter 2 of [30] and [8] for details.

The analysis performed in Section 5 below, based on a construction of a Lyapunov functional, gives the sufficient condition

$$2(1 + \lambda)\tau < 1 \quad (15)$$

for asymptotic stability of the trivial steady state. This is clearly not optimal, since we know that for $\lambda = 0$ the sufficient and necessary condition for asymptotic stability is $2\tau < \pi/2$. Consequently, to get a more complete picture, we studied the asymptotic behavior of (13) numerically. Again, we discretized (13) on an equidistant mesh with time step $\Delta t > 0$, chosen such that τ is an integer multiple of Δt . Implicit Euler discretization of \dot{x} and $\dot{\tilde{x}}$ leads then to a difference equation that can be numerically integrated. We chose the constant initial datum $x^0(t) \equiv 1$ for $t \in [-\tau, 0]$ and performed systematic simulations with varying parameter values $\lambda \geq 0$ and $\tau \geq 0$, detecting either decay or growth of the amplitude of the oscillations of the solution. The result is presented in Fig. 3, where the solid curve represents the threshold for stability of the zero steady state obtained numerically – the values of (λ, τ) below the solid curve give $x(t) \rightarrow 0$ for large $t > 0$, while the values above the curve produce diverging $x(t)$. For comparison, the dashed curve represents the analytical sufficient condition (15). The critical value $\tau = \pi/4$ for the system without anticipation is indicated by the dotted line.

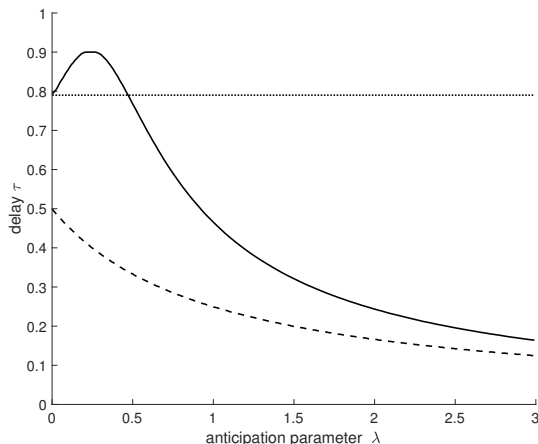


Figure 3: Systematic simulations of (13) subject to the constant initial datum $x^0(t) \equiv 1$ for $t \in [-\tau, 0]$, for the parameter range $\lambda \in [0, 3]$ and $\tau \in [0, 1]$. The solid curve represents the threshold for stability of the zero steady state obtained numerically – the values of (λ, τ) below the solid curve give $x(t) \rightarrow 0$ for large $t > 0$, while the values above the curve produce diverging $x(t)$. The dashed curve represents the analytical sufficient condition (15). The critical value $\tau = \pi/4$ for the system without anticipation is indicated by the dotted line.

The most interesting set of parameter values in Fig. 3 is located in the area enclosed between the solid curve and the dotted line $\tau = \pi/4$. Here the zero steady state is stable with anticipation, but becomes unstable when anticipation is suppressed (i.e., $\lambda = 0$). This stabilizing effect of anticipation only works for values of τ below approx. 0.9, with the range of admissible values of λ getting narrower as τ approaches this critical value. To illustrate this stabilizing effect, we plotted numerical solutions of (13) for $\tau = 0.85$ and $\lambda \in \{0.04, 0.25, 0.45\}$ represented by the solid line in Fig. 4, while, for comparison, the dashed line shows the solution without anticipation (i.e., $\lambda = 0$). We see that for $\lambda = 0.02$ the effect of anticipation significantly reduces the amplitude of oscillations of the solution, but still is too weak to ultimately stabilize it. The choice $\lambda = 0.25$ gives stability of the zero steady state, but increasing λ to 0.45 makes it unstable again. Unfortunately, we are not able to provide analytical insights into this phenomenon. In particular, established methods for asymptotic analysis of functional differential equations (Lyapunov-type functionals, characteristic polynomials) deliver sufficient/necessary conditions for stability of the form of smallness assumptions on the parameters; see (11). However, in Figs. 3 and 4 we observe stabilization for moderate levels of anticipation, where the value of λ is neither too small nor too large. We therefore hypothesize that this effect is not explainable by established analytical methods and new approaches need to be developed. We leave it as subject of future work.

4 Asymptotic consensus with transmission delay - proof of Theorem 1

In this Section we provide a proof of Theorem 1. It is based on the following geometrical lemma, presented in [13], based on the notion of coefficient of ergodicity introduced by Dobrushin [7] in the context of Markov chains.

Lemma 1. *Let $N \geq 2$ and $\{x_1, \dots, x_N\} \subset \mathbb{R}^d$ be any set of vectors in \mathbb{R}^d . Fix $i, k \in [N]$ such that $i \neq k$*

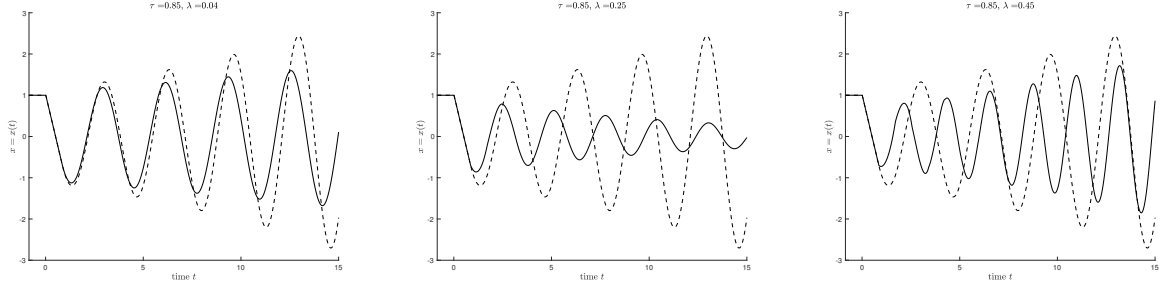


Figure 4: Numerical integration of (13) subject to the constant initial datum $x^0 \equiv 0$, with $\tau = 0.85$ and $\lambda = 0.04$ (solid line in left panel), $\lambda = 0.25$ (solid line in middle panel) and $\lambda = 0.45$ (solid line in right panel). For comparison, numerical solution of (13) with $\tau = 0.85$ and $\lambda = 0$ (no anticipation) is indicated by dashed line in all three panels.

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

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

Denote

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

Then

$$\left| \sum_{j \neq i} \eta_j^i x_j - \sum_{j \neq k} \eta_j^k x_j \right| \leq (1 - (N - 2)\mu) \max_{i, j \in [N]} |x_i - x_j|.$$

We refer to [13, Lemma 3.2] for the proof. Note that, clearly, Lemma 1 only provides a nontrivial bound if $N \geq 3$; however, the case $N = 2$ was already discussed in Section 3.1.

Let us recall that due to the positivity assumption (8), all off-diagonal elements ψ_{ij} , $i \neq j$, are strictly positive. Therefore, introducing the notation

$$\underline{\psi} := \min_{i \neq j} \psi_{ij},$$

we have $\underline{\psi} > 0$. We further denote

$$\gamma := 1 - (N - 2)\underline{\psi}. \tag{16}$$

Due to the stochasticity assumption (7) we have

$$1 = \sum_{j \neq i} \psi_{ij} \geq (N - 1)\underline{\psi},$$

so that

$$\gamma \geq 1 - \frac{N - 2}{N - 1} > 0.$$

Let us also recall the assumption that there exists some $T_0 > 0$ and fixed indices $I, J \in [N]$ such that (9) holds for $t \geq T_0$. We then define the functional

$$\mathcal{L}(t) := \frac{1}{2} \left[(x_I - \lambda\tau \tilde{\Psi}^I) - (x_K - \lambda\tau \tilde{\Psi}^K) \right]^2 + \left[\frac{1 + \lambda\tau}{2\gamma} - \lambda\tau \right] \gamma^2 \int_{t-\tau}^t d_x(s)^2 ds, \tag{17}$$

where we introduced the short-hand notation

$$\Psi^i(t) := \sum_{j \neq i} \psi_{ij} x_j(t), \quad (18)$$

for $i \in [N]$. Note that we again apply the convention $\tilde{\Psi}^i = \Psi^i(t - \tau)$. Observe that since $\gamma \in (0, 1]$, the term $\left[\frac{1+\lambda\tau}{2\gamma} - \lambda\tau\right]$ is nonnegative whenever $\lambda\tau \leq 1$, and then $\mathcal{L} \geq 0$ by definition. We now derive a decay estimate on $\mathcal{L} = \mathcal{L}(t)$.

Lemma 2. *Let the matrix of communication weights $(\psi_{ij})_{i,j \in [N]}$ verify the stochasticity (7) and positivity (8) assumptions. Let $\lambda\tau \leq 1$ and $\gamma \in (0, 1]$ be given by (16). Then for $\mathcal{L} = \mathcal{L}(t)$ given by (17) we have along solutions of (2), for $t > T_0$,*

$$\frac{d}{dt} \mathcal{L}(t) \leq -(1 - \gamma)(1 - \lambda\tau\gamma) d_x(t)^2. \quad (19)$$

Proof. We have for all $t > T_0$,

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \left[\left(x_I - \lambda\tau \tilde{\Psi}^I \right) - \left(x_K - \lambda\tau \tilde{\Psi}^K \right) \right]^2 &= - \left(x_I - x_K - \lambda\tau \left(\tilde{\Psi}^I - \tilde{\Psi}^K \right) \right) \cdot \left(x_I - x_K - \left(\tilde{\Psi}^I - \tilde{\Psi}^K \right) \right) \\ &= -|x_I - x_K|^2 - \lambda\tau \left| \tilde{\Psi}^I - \tilde{\Psi}^K \right|^2 + (1 + \lambda\tau)(x_I - x_K) \cdot \left(\tilde{\Psi}^I - \tilde{\Psi}^K \right). \end{aligned}$$

We apply the Cauchy-Schwarz inequality with $\delta > 0$ for the last term,

$$(1 + \lambda\tau)(x_I - x_K) \cdot \left(\tilde{\Psi}^I - \tilde{\Psi}^K \right) \leq \frac{(1 + \lambda\tau)\delta}{2} |x_I - x_K|^2 + \frac{1 + \lambda\tau}{2\delta} \left| \tilde{\Psi}^I - \tilde{\Psi}^K \right|^2,$$

which gives, using (9),

$$\frac{1}{2} \frac{d}{dt} \left[\left(x_I - \lambda\tau \tilde{\Psi}^I \right) - \left(x_K - \lambda\tau \tilde{\Psi}^K \right) \right]^2 \leq \left[\frac{(1 + \lambda\tau)\delta}{2} - 1 \right] d_x(t)^2 + \left[\frac{1 + \lambda\tau}{2\delta} - \lambda\tau \right] \left| \tilde{\Psi}^I - \tilde{\Psi}^K \right|^2.$$

With Lemma 1, recalling (16), we have

$$\left| \tilde{\Psi}^I - \tilde{\Psi}^K \right| = \left| \sum_{j \neq i} \psi_{ij} \tilde{x}_j - \sum_{j \neq k} \psi_{kj} \tilde{x}_j \right| \leq (1 - (N - 2)\underline{\psi}) \max_{i,j \in [N]} |\tilde{x}_I - \tilde{x}_j| = \gamma d_x(t - \tau). \quad (20)$$

Consequently, making the assumption

$$\delta \leq \frac{1 + \lambda\tau}{2\lambda\tau}, \quad (21)$$

so that $\left[\frac{1+\lambda\tau}{2\delta} - \lambda\tau\right] \geq 0$, we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \left[\left(x_I - \lambda\tau \tilde{\Psi}^I \right) - \left(x_K - \lambda\tau \tilde{\Psi}^K \right) \right]^2 &\leq \left[\frac{(1 + \lambda\tau)\delta}{2} - 1 \right] d_x(t)^2 \\ &\quad + \left[\frac{1 + \lambda\tau}{2\delta} - \lambda\tau \right] \gamma^2 d_x(t - \tau)^2. \end{aligned}$$

Therefore, defining

$$\mathcal{L}_\delta(t) := \frac{1}{2} \left[\left(x_I - \lambda\tau \tilde{\Psi}^I \right) - \left(x_K - \lambda\tau \tilde{\Psi}^K \right) \right]^2 + \left[\frac{1 + \lambda\tau}{2\delta} - \lambda\tau \right] \gamma^2 \int_{t-\tau}^t d_x(s)^2 ds,$$

we have

$$\frac{d}{dt} \mathcal{L}_\delta(t) \leq \left(\left[\frac{(1 + \lambda\tau)\delta}{2} - 1 \right] + \left[\frac{1 + \lambda\tau}{2\delta} - \lambda\tau \right] \gamma^2 \right) d_x(t)^2.$$

Minimization of the right-hand side in $\delta > 0$ motivates the choice $\delta := \gamma$, which identifies \mathcal{L}_δ with \mathcal{L} defined in (17), and

$$\frac{d}{dt} \mathcal{L}(t) \leq -(1 - \gamma)(1 - \lambda\tau\gamma) d_x(t)^2.$$

We finally note that with $\delta := \gamma$ and $\lambda\tau \leq 1$, assumption (21) is indeed verified since its right-hand side is larger or equal 1 while $\delta = \gamma \leq 1$. \square

Obviously, for $\lambda\tau \leq 1$ and $\gamma \in (0, 1)$, Lemma 2 states that \mathcal{L} is a monotonically decaying functional, and the decay is strict unless $d_x \equiv 0$. However, this per se does not imply that $d_x(t) \rightarrow 0$ as $t \rightarrow \infty$. Moreover, we cannot apply the Lyapunov theorem here (strictly speaking, its version adapted to neutral functional differential equations, see [9, Theorem 8.1 of Section 9]), since system (2) has a continuum of steady states characterized by $x_i \equiv x_j$ for all $i, j \in [N]$, and the asymptotic one is not determined by the initial datum due to the lack of conserved quantities. Consequently, to conclude the convergence to consensus, we shall employ Barbalat's lemma [1]. For its application we need uniform continuity of $d_x = d_x(t)$, which is a direct consequence on the following uniform a-priori estimate on the derivative of the trajectories $x_i = x_i(t)$.

Lemma 3. *Let $\lambda\tau \leq 1$ and denote*

$$M := \max_{t \in [-\tau, 0]} \max_{i \in [N]} \{|x_i^0(t)|, |\dot{x}_i^0(t)|\}. \quad (22)$$

Then for all $t > 0$ and $i \in [N]$,

$$|x_i(t)| \leq (1 + \lambda\tau)M, \quad |\dot{x}_i(t)| \leq 3(1 + \lambda\tau)M. \quad (23)$$

Proof. Let us denote, for $k \in \mathbb{N}$,

$$y_i^k(t) := x_i(t + (k - 1)\tau) \quad \text{for } t \in [0, \tau], \quad i \in [N].$$

Then (2) is rewritten in terms of y_i^k as

$$\dot{y}_i^{k+1} = \sum_{j=1}^N \psi_{ij} (y_j^k + \lambda\tau \dot{y}_j^k - y_i^{k+1}).$$

A simple induction argument shows that for all $k \in \mathbb{N}$ and $i \in [N]$,

$$\begin{aligned} \dot{y}_i^{k+1} = (1 - \lambda\tau) & \left[\sum_{j_k=1}^N \psi_{i,j_k} y_{j_k}^k + \lambda\tau \sum_{j_k=1}^N \sum_{j_{k-1}=1}^N \psi_{i,j_k} \psi_{j_k,j_{k-1}} y_{j_{k-1}}^{k-1} + \dots \right. \\ & \left. \dots + (\lambda\tau)^{k-1} \sum_{j_k=1}^N \dots \sum_{j_1=1}^N \psi_{i,j_k} \dots \psi_{j_2,j_1} y_{j_1}^1 \right] \\ & + (\lambda\tau)^k \sum_{j_k=1}^N \dots \sum_{j_0=1}^N \psi_{i,j_k} \dots \psi_{j_1,j_0} y_{j_0}^0 \\ & + (\lambda\tau)^{k+1} \sum_{j_k=1}^N \dots \sum_{j_0=1}^N \psi_{i,j_k} \dots \psi_{j_1,j_0} y_{j_0}^0 - y_i^{k+1}. \end{aligned} \quad (24)$$

We shall first prove the first claim of (23), i.e., that for all $k \in \mathbb{N}$,

$$|y_i^k(t)| \leq (1 + \lambda\tau)M \quad \text{for all } t \in [0, \tau], \quad i \in [N]. \quad (25)$$

Obviously, $y_i^0 \equiv x_i^0$ for all $i \in [N]$, so that (25) holds for $k = 0$ due to (22). We proceed inductively, assuming that (25) holds for $k = 0, \dots, K$. With the Cauchy-Schwarz inequality, using $\lambda\tau \leq 1$ and (7), we infer from (24),

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} |y_i^{K+1}(t)|^2 &\leq \left((1 - \lambda\tau) \sum_{k=0}^{K-1} (\lambda\tau)^k + (\lambda\tau)^K + (\lambda\tau)^{K+1} \right) M |y_i^{K+1}(t)| - |y_i^{K+1}(t)|^2 \\ &= \left[(1 + (\lambda\tau)^{K+1}) M - |y_i^{K+1}(t)| \right] |y_i^{K+1}(t)| \end{aligned}$$

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

$$\frac{d}{dt} |y_i^{K+1}(t)| \leq (1 + (\lambda\tau)^{K+1}) M - |y_i^{K+1}(t)|. \quad (26)$$

Then, integration in time gives

$$\begin{aligned} |y_i^{K+1}(t)| &\leq e^{-t} |y_i^{K+1}(0)| + (1 - e^{-t}) (1 + (\lambda\tau)^{K+1}) M \\ &\leq e^{-t} (1 + \lambda\tau)M + (1 - e^{-t}) (1 + (\lambda\tau)^{K+1}) M \\ &\leq (1 + \lambda\tau)M, \end{aligned}$$

for $t \in [0, \tau]$, where we used the induction hypothesis $|y_i^{K+1}(0)| = |y_i^K(\tau)| \leq (1 + \lambda\tau)M$ for the second inequality and $\lambda\tau \leq 1$ for the third one.

Finally, using (25) in (24) gives

$$\begin{aligned} |y_i^{K+1}| &\leq \left((1 - \lambda\tau) \sum_{k=0}^{K-1} (\lambda\tau)^k + (\lambda\tau)^K + (\lambda\tau)^{K+1} + 1 \right) (1 + \lambda\tau)M \\ &= (2 + (\lambda\tau)^{K+1}) (1 + \lambda\tau)M \end{aligned}$$

for all $K \in \mathbb{N}$, which with $\lambda\tau \leq 1$ implies the second claim of (23). \square

We now provide the proof of Theorem 1.

Proof. An integration of (19) on the interval (T_0, t) gives

$$(1 - \gamma)(1 - \lambda\tau\gamma) \int_{T_0}^t d_x(s)^2 ds \leq \mathcal{L}(T_0) - \mathcal{L}(t) \leq \mathcal{L}(T_0).$$

Therefore, since for $N \geq 3$ we have $\gamma < 1$ by (16), the squared diameter $d_x^2 = d_x^2(t)$ must be integrable as $t \rightarrow \infty$. Moreover, with (23) we have

$$\left| \frac{d}{dt} d_x(t)^2 \right| \leq 2|x_i - x_k| |\dot{x}_i - \dot{x}_k| \leq 24(1 + \lambda\tau)^2 M^2.$$

Then, an application of the Barbalat lemma [1] implies that $\lim_{t \rightarrow \infty} d_x(t) = 0$. \square

Remark 1. Let us now explain why we had to adopt the assumption (9) of Theorem 1. Observe that, due to the continuity of the trajectories $x_i = x_i(t)$, there is an at most countable system of open, mutually disjoint intervals $\{\mathcal{I}_\sigma\}_{\sigma \in \mathbb{N}}$ such that

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

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

$$d_x(t) = |x_{i(\sigma)}(t) - x_{k(\sigma)}(t)| \quad \text{for } t \in \mathcal{I}_\sigma. \quad (27)$$

Obviously, assumption (9) means that the system $\{\mathcal{I}_\sigma\}_\sigma$ can be chosen of finite cardinality, with the last interval being (T_0, ∞) .

If assumption (9) was not met, i.e., the system $\{\mathcal{I}_\sigma\}_{\sigma \in \mathbb{N}}$ was countably infinite, we would have to replace the functional (17) by

$$\mathcal{L}'(t) := \frac{1}{2} \left[\left(x_{i(\sigma)} - \lambda\tau \tilde{\Psi}^{i(\sigma)} \right) - \left(x_{k(\sigma)} - \lambda\tau \tilde{\Psi}^{k(\sigma)} \right) \right]^2 + \left[\frac{1 + \lambda\tau}{2\gamma} - \lambda\tau \right] \gamma^2 \int_{t-\tau}^t d_x(s)^2 ds,$$

with $i = i(\sigma), k = k(\sigma)$ given by (27). Then, a minimal adaptation of the proof of Lemma 2 would give the dissipation estimate (19) for $\mathcal{L}' = \mathcal{L}'(t)$, for all $t \in \bigcup_{\sigma \in \mathbb{N}} \mathcal{I}_\sigma$, i.e., for almost all $t > 0$. However, the functional $\mathcal{L}'(t)$ is not globally continuous since the terms $\tilde{\Psi}^{i(\sigma)}, \tilde{\Psi}^{k(\sigma)}$ may have jumps across the boundaries of the intervals \mathcal{I}_σ . Although the jumps can be controlled in spirit of (20), this control is not sufficient to finally conclude that $\lim_{t \rightarrow \infty} d_x(t) = 0$. This is the reason why we had to admit the assumption (9) in the formulation of Theorem 1.

5 Asymptotic consensus with reaction delay - proof of Theorem 2

Let us recall that Theorem 2 assumes the matrix of communication weights $(\psi_{ij})_{i,j \in [N]}$ to be symmetric and irreducible. Since the values of the diagonal entries ψ_{ii} are irrelevant for the dynamics of (3), we can formally set them such that all row and column sums are the same. In particular, without loss of generality (up to an eventual rescaling of time), we shall assume that $(\psi_{ij})_{i,j \in [N]}$ is a bi-stochastic matrix, i.e.,

$$\sum_{j=1}^N \psi_{ij} = \sum_{j=1}^N \psi_{ji} = 1 \quad \text{for all } i \in [N], \quad (28)$$

The symmetry implies that the mean opinion $X = X(t)$ defined in (5) is conserved along the solutions of (3). Therefore, without loss of generality, we shall assume that $X(t) \equiv 0$ for all $t \geq 0$.

For $t \geq 0$ we define the quantity

$$D(t) := \sum_{i=1}^N \sum_{j=1}^N \psi_{ij} |x_j - x_i|^2, \quad (29)$$

with the convention $\tilde{D} := D(t - \tau) = \sum_{i=1}^N \sum_{j=1}^N \psi_{ij} |\tilde{x}_j - \tilde{x}_i|^2$. We also introduce the short-hand notation

$$\Phi^i := \sum_{j=1}^N \psi_{ij} (x_j - x_i), \quad (30)$$

and, again, $\tilde{\Phi}^i = \Phi^i(t - \tau)$.

Lemma 4. *For any $\varepsilon > 0$ and $\delta > 0$ we have, along the solutions of (3) with bi-stochastic interaction weights (28),*

$$\begin{aligned} \sum_{i=1}^N \tilde{\Phi}^i \cdot x_i &\leq \frac{\delta - 1}{2} D(t - \tau) + \frac{(1 + \varepsilon)\tau}{2\delta} \int_{t-\tau}^t D(s - \tau) ds \\ &\quad + \frac{\lambda^2 \tau^2}{\delta} (1 + \varepsilon^{-1}) \sum_{i=1}^N \left(|\Phi^i(t - \tau)|^2 + |\Phi^i(t - 2\tau)|^2 \right), \end{aligned} \quad (31)$$

for all $t > \tau$.

Proof. With (3) we have

$$\begin{aligned}\sum_{i=1}^N \tilde{\Phi}^i \cdot x_i &= \sum_{i=1}^N \sum_{j=1}^N \psi_{ij}(\tilde{x}_j - \tilde{x}_i) \cdot x_i \\ &= \sum_{i=1}^N \sum_{j=1}^N \psi_{ij}(\tilde{x}_j - \tilde{x}_i) \cdot \tilde{x}_i + \sum_{i=1}^N \sum_{j=1}^N \psi_{ij}(\tilde{x}_j - \tilde{x}_i) \cdot (x_i - \tilde{x}_i).\end{aligned}$$

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

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

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

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

where we used (28) in the second line. Hence,

$$\sum_{i=1}^N \tilde{\Phi}^i \cdot x_i \leq \frac{\delta - 1}{2} D(t - \tau) + \frac{1}{2\delta} \sum_{i=1}^N |x_i - \tilde{x}_i|^2. \quad (32)$$

Next, for $t > \tau$ we use (3) to evaluate the difference $x_i - \tilde{x}_i$,

$$\begin{aligned}x_i - \tilde{x}_i &= \int_{t-\tau}^t \dot{x}_i(s) \, ds \\ &= \int_{t-\tau}^t \Phi^i(s - \tau) + \lambda \tau \dot{\Phi}^i(s - \tau) \, ds \\ &= \int_{t-\tau}^t \Phi^i(s - \tau) \, ds + \lambda \tau (\Phi^i(t - \tau) - \Phi^i(t - 2\tau)).\end{aligned}$$

Taking the square, using the inequality $(a + b)^2 \leq (1 + \varepsilon)a^2 + (1 + \varepsilon^{-1})b^2$ with $\varepsilon > 0$, and summing over $i \in [N]$ gives

$$\begin{aligned}\frac{1}{2\delta} \sum_{i=1}^N |x_i - \tilde{x}_i|^2 &\leq \frac{1 + \varepsilon}{2\delta} \sum_{i=1}^N \left| \int_{t-\tau}^t \Phi^i(s - \tau) \, ds \right|^2 + \frac{\lambda^2 \tau^2}{2\delta} (1 + \varepsilon^{-1}) \sum_{i=1}^N |\Phi^i(t - \tau) - \Phi^i(t - 2\tau)|^2 \\ &\leq \frac{(1 + \varepsilon)\tau}{2\delta} \sum_{i=1}^N \int_{t-\tau}^t |\Phi^i(s - \tau)|^2 \, ds + \frac{\lambda^2 \tau^2}{\delta} (1 + \varepsilon^{-1}) \sum_{i=1}^N (|\Phi^i(t - \tau)|^2 + |\Phi^i(t - 2\tau)|^2),\end{aligned}$$

where we used the Cauchy-Schwarz inequality in the second line. The first term of the right-hand side is

estimated using the discrete Jensen inequality, recalling the stochasticity (28),

$$\begin{aligned}
\frac{(1+\varepsilon)\tau}{2\delta} \sum_{i=1}^N \int_{t-\tau}^t |\Phi^i(s-\tau)|^2 ds &= \frac{(1+\varepsilon)\tau}{2\delta} \sum_{i=1}^N \int_{t-\tau}^t \left| \sum_{j=1}^N \psi_{ij}(x_j(s-\tau) - x_i(s-\tau)) \right|^2 ds \\
&\leq \frac{(1+\varepsilon)\tau}{2\delta} \sum_{i=1}^N \int_{t-\tau}^t \sum_{j=1}^N \psi_{ij} |x_j(s-\tau) - x_i(s-\tau)|^2 ds \\
&= \frac{(1+\varepsilon)\tau}{2\delta} \int_{t-\tau}^t D(s-\tau) ds.
\end{aligned}$$

Consequently, we arrive at

$$\frac{1}{2\delta} \sum_{i=1}^N |x_i - \tilde{x}_i|^2 \leq \frac{(1+\varepsilon)\tau}{2\delta} \int_{t-\tau}^t D(s-\tau) ds + \frac{\lambda^2\tau^2}{\delta} (1+\varepsilon^{-1}) \sum_{i=1}^N \left(|\Phi^i(t-\tau)|^2 + |\Phi^i(t-2\tau)|^2 \right),$$

and inserting into (32) gives (31). \square

For $\varepsilon > 0$ and $\delta > 0$ to be specified later, we define the functional

$$\begin{aligned}
\mathcal{L}_{\varepsilon,\delta}(t) &:= \frac{1}{2} \sum_{i=1}^N \left(x_i - \lambda\tau\tilde{\Phi}^i \right)^2 + \frac{\lambda^2\tau^2}{\delta} (1+\varepsilon^{-1}) \int_{t-\tau}^t \sum_{i=1}^N |\Phi^i(s-\tau)|^2 ds \\
&\quad + \frac{(1+\varepsilon)\tau}{2\delta} \int_{t-\tau}^t \int_{\theta}^t D(s-\tau) ds d\theta.
\end{aligned} \tag{33}$$

Lemma 5. *For $\varepsilon := 2\lambda$ and $\delta := (1+2\lambda)\tau$, we have along the solutions of (3) with bi-stochastic interaction weights (28),*

$$\frac{d}{dt} \mathcal{L}_{\varepsilon,\delta}(t) \leq \left((1+\lambda)\tau - \frac{1}{2} \right) D(t-\tau), \tag{34}$$

for all $t > \tau$, with $\mathcal{L}_{\varepsilon,\delta} = \mathcal{L}_{\varepsilon,\delta}(t)$ defined in (33).

Proof. Using (3), we have for all $i \in [N]$,

$$\frac{1}{2} \frac{d}{dt} \left(x_i - \lambda\tau\tilde{\Phi}^i \right)^2 = \left(x_i - \lambda\tau\tilde{\Phi}^i \right) \cdot \tilde{\Phi}^i = -\lambda\tau \left| \tilde{\Phi}^i \right|^2 + \tilde{\Phi}^i \cdot x_i.$$

Summing over $i \in [N]$ and using (31), recalling the notation $\tilde{\Phi}^i = \Phi^i(t-\tau)$,

$$\begin{aligned}
\frac{1}{2} \frac{d}{dt} \sum_{i=1}^N \left(x_i - \lambda\tau\tilde{\Phi}^i \right)^2 &\leq \frac{\delta-1}{2} D(t-\tau) + \frac{(1+\varepsilon)\tau}{2\delta} \int_{t-\tau}^t D(s-\tau) ds \\
&\quad + \left(\frac{\lambda^2\tau^2}{\delta} (1+\varepsilon^{-1}) - \lambda\tau \right) \sum_{i=1}^N |\Phi^i(t-\tau)|^2 \\
&\quad + \frac{\lambda^2\tau^2}{\delta} (1+\varepsilon^{-1}) \sum_{i=1}^N |\Phi^i(t-2\tau)|^2.
\end{aligned}$$

Noting that

$$\frac{\lambda^2\tau^2}{\delta} (1+\varepsilon^{-1}) \frac{d}{dt} \int_{t-\tau}^t \sum_{i=1}^N |\Phi^i(s-\tau)|^2 ds = \frac{\lambda^2\tau^2}{\delta} (1+\varepsilon^{-1}) \sum_{i=1}^N \left(|\Phi^i(t-\tau)|^2 - |\Phi^i(t-2\tau)|^2 \right),$$

and

$$\frac{(1+\varepsilon)\tau}{2\delta} \frac{d}{dt} \int_{t-\tau}^t \int_{\theta}^t D(s-\tau) ds d\theta = \frac{(1+\varepsilon)\tau^2}{2\delta} D(t-\tau) - \frac{(1+\varepsilon)\tau}{2\delta} \int_{t-\tau}^t D(s-\tau) ds,$$

we have

$$\frac{d}{dt} \mathcal{L}_{\varepsilon, \delta}(t) \leq \left(\frac{\delta-1}{2} + \frac{(1+\varepsilon)\tau^2}{2\delta} \right) D(t-\tau) + \left(\frac{2\lambda^2\tau^2}{\delta} (1+\varepsilon^{-1}) - \lambda\tau \right) \sum_{i=1}^N |\Phi^i(t-\tau)|^2.$$

Next, the discrete Jensen inequality, recalling the stochasticity (28), gives

$$\sum_{i=1}^N |\Phi^i(t-\tau)|^2 = \sum_{i=1}^N \left| \sum_{j=1}^N \psi_{ij} (\tilde{x}_j - \tilde{x}_i) \right|^2 \leq \sum_{i=1}^N \sum_{j=1}^N \psi_{ij} |\tilde{x}_j - \tilde{x}_i|^2 = D(t-\tau).$$

Consequently, assuming that λ , τ and ε , δ are such that

$$\frac{2\lambda^2\tau^2}{\delta} (1+\varepsilon^{-1}) - \lambda\tau \geq 0, \quad (35)$$

we have

$$\frac{d}{dt} \mathcal{L}_{\varepsilon, \delta}(t) \leq \left(\frac{\delta-1}{2} + \frac{(1+\varepsilon)\tau^2}{2\delta} + \frac{2\lambda^2\tau^2}{\delta} (1+\varepsilon^{-1}) - \lambda\tau \right) D(t-\tau).$$

We first minimize the right-hand side in $\varepsilon > 0$, which leads to $\varepsilon := 2\lambda$ and

$$\frac{d}{dt} \mathcal{L}_{\varepsilon, \delta}(t) \leq \left(\frac{\delta-1}{2} + \frac{(1+2\lambda)^2\tau^2}{2\delta} - \lambda\tau \right) D(t-\tau).$$

Now we minimize the right-hand side in $\delta > 0$. This gives the optimal choice $\delta := (1+2\lambda)\tau$ and

$$\frac{d}{dt} \mathcal{L}_{\varepsilon, \delta}(t) \leq \left((1+\lambda)\tau - \frac{1}{2} \right) D(t-\tau),$$

which is (34). It remains to check that condition (35) is verified. Remarkably, a simple calculation shows that with $\varepsilon := 2\lambda$ and $\delta := (1+2\lambda)\tau$, (35) holds with equality. \square

We are now in position to present the proof of Theorem 2.

Proof. We first note that, assuming $X = 0$, the only stationary solution of (3) is the trivial solution $x_i = 0$ for all $i \in [N]$. Indeed, the stationary version of (3) can be written as

$$-L\mathbf{x} = 0, \quad (36)$$

where $\mathbf{x} = (x_1, \dots, x_N)$ and $L \in \mathbb{R}^{N \times N}$ is the discrete Laplacian

$$L_{ij} := -\psi_{ij} \quad \text{for } i \neq j, \quad L_{ii} := \sum_{j \neq i} \psi_{ij} \quad \text{for } i \in [N].$$

Since $(\psi_{ij})_{i,j \in [N]}$ is assumed to be irreducible, basic theory [21] states that the null space of L is the diagonal $\{(x, \dots, x) \in \mathbb{R}^{Nd}; x \in \mathbb{R}^d\}$. But then, the assumption $X = 0$ implies that (36) holds if and only if $\mathbf{x} = 0$.

Consequently, global asymptotic consensus for (3) is equivalent to the asymptotic stability of the trivial steady state $x_i = 0$ for all $i \in [N]$. But the asymptotic stability is implied by the fact that, if (11) holds,

then, by Lemma 5, $\mathcal{L}_{\varepsilon, \delta}$ with $\varepsilon := 2\lambda$ and $\delta := (1 + 2\lambda)\tau$ is a Lyapunov functional for (3). Indeed, due to the irreducibility of $(\psi_{ij})_{i, j \in [N]}$,

$$D(t) = \sum_{i=1}^N \sum_{j=1}^N \psi_{ij} |x_j - x_i|^2 = L\mathbf{x} \cdot \mathbf{x}$$

vanishes if and only if $\mathbf{x} = 0$. Consequently, the right-hand side of (34) is nonpositive and zero only for the trivial steady state. An application of the Lyapunov theorem for systems of neutral functional differential equations, see, e.g., [9, Theorem 8.1 of Section 9], provides then the sought global asymptotic stability of the trivial steady state $x_i = 0$ for all $i \in [N]$. \square

Acknowledgment

JH acknowledges the support of the KAUST baseline funds.

References

- [1] I. Barbalat: *Systèmes d'équations différentielles d'oscillations nonlinéaires*. Rev. Math. Pures Appl. 4, 2 (1959), 267–270.
- [2] 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.
- [3] C. Castellano, S. Fortunato and V. Loreto: *Statistical physics of social dynamics*. Rev. Mod. Phys., 81, (2009), 591–646.
- [4] 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.
- [5] F. Cucker and S. Smale, Emergent behaviour in flocks, *IEEE T. on Automat. Contr.*, **52** (2007), 852–862.
- [6] F. Cucker and S. Smale, On the mathematics of emergence, *Jap. J. Math.*, **2** (2007), 197–227.
- [7] R. L. Dobrushin, *Central limit theorem for non-stationary Markov chains. I*, Theory Probab. Appl., 1:1 (1956), 65–80.
- [8] GYORI I., AND LADAS G. *Oscillation Theory of Delay Differential Equations with Applications*. Oxford Science Publications, Clarendon Press, Oxford, 1991.
- [9] J. Hale and S. Verduyn Lunel: *Introduction to Functional Differential Equations*. Applied Mathematical Sciences Vol. 99, Springer Science+Business Media, 1993.
- [10] H. Hamman: *Swarm Robotics: A Formal Approach*. Springer, 2018.
- [11] J. Haskovec: *Asymptotic consensus in the Hegselmann-Krause model with finite speed of information propagation*. Proc. Amer. Math. Soc., to appear (2021).
- [12] J. Haskovec: *Direct proof of unconditional asymptotic consensus in the Hegselmann-Krause model with transmission-type delay*. Bull. London Math. Soc., to appear (2021).
- [13] 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), 130–148.

- [14] R. Hegselmann and U. Krause, *Opinion dynamics and bounded confidence models, analysis, and simulation*, J. Artif. Soc. Soc. Simul., 5, (2002), 1–24.
- [15] P.E. Jabin and S. Motsch: *Clustering and asymptotic behavior in opinion formation*. J. Differential Equations 257 (2014), 4165–4187.
- [16] 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.
- [17] B. Kim, J. Kwon, S. Choi, and J. Yand: *Feedback Stabilization of First Order Neutral Delay Systems Using the Lambert W Function*. Appl. Sci. 9 (2019), 3539.
- [18] P. Krugman: *The Self Organizing Economy*. Blackwell Publishers, 1995.
- [19] 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.
- [20] 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.
- [21] B. Mohar: *The Laplacian spectrum of graphs*. In Y. Alavi, G. Chartrand, O.R. Oellermann, and A.J. Schwenk (eds.), Graph Theory, Combinatorics, and Applications, volume 2, pp. 871–898. John Wiley & Sons, 1991.
- [22] 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.
- [23] S.-I. Niculescu: *Delay Effects on Stability. A Robust Control Approach*. Springer-Verlag London, 2001.
- [24] A. Paolucci: *Convergence to consensus for a Hegselmann-Krause-type model with distributed time delay*. Preprint (2020), arxiv.org/abs/2005.08030.
- [25] 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.
- [26] 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.
- [27] R. Shu and E. Tadmor: *Anticipation Breeds Alignment*. Arch. Rational Mech. Anal. 240 (2021), 203–241.
- [28] 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.
- [29] 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).
- [30] H. Smith: *An Introduction to Delay Differential Equations with Applications to the Life Sciences*. Springer New York Dordrecht Heidelberg London, 2011.
- [31] K. Szwajkowska, 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.
- [32] 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.
- [33] T. Vicsek and A. Zafeiris: *Collective motion*. Phys. Rep., 517 (2012), 71–140.
- [34] 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.