

Stability by Fixed Point Theory in Time Delayed Distributed Consensus Networks

Christoforos Somarakis, John S. Baras, and Evrripides Paraskevas

The
Institute for
Systems
Research



A. JAMES CLARK
SCHOOL OF ENGINEERING

ISR develops, applies and teaches advanced methodologies of design and analysis to solve complex, hierarchical, heterogeneous and dynamic problems of engineering technology and systems for industry and government.

ISR is a permanent institute of the University of Maryland, within the A. James Clark School of Engineering. It is a graduated National Science Foundation Engineering Research Center.

www.isr.umd.edu

STABILITY BY FIXED POINT THEORY IN TIME DELAYED DISTRIBUTED CONSENSUS NETWORKS

CHRISTOFOROS SOMARAKIS¹, JOHN S. BARAS² AND EVRIPIDIS PARASKEVAS³

The Institute For Systems Research
University Of Maryland
College Park, MD 20740, USA

ABSTRACT. We study the stability of linear time invariant distributed consensus dynamics in the presence of multiple propagation and processing delays. We employ fixed point theory (FPT) methods and derive sufficient conditions for asymptotic convergence to a common value while the emphasis is given in estimating the rate of convergence. We argue that this approach is novel in the field of networked dynamics as it is also flexible and thus capable of analyzing a wide variety of consensus based algorithms for which conventional Lyapunov methods are either too restrictive or unsuccessful.

1. Introduction. Distributed dynamics have, over the past decade, carried the beacon of research in the control community. Starting from the seminal work of Tsitsiklis [27] the subject was reheated with the work of Jadbabaie et al. [7] who gave a rigorous proof of the leaderless co-ordination in a flocking model proposed by Viscek et al. [28].

Since then, an enormous amount of research has been produced from different fields of Applied Science concerning types of coordination among autonomous agents who exchange information with application in co-operation, formation control, co-ordination (e.g. deterministic or stochastic) and various communication conditions (see for example [1, 11, 10, 4, 17, 9, 16, 14, 15, 12] and references therein).

All of the proposed models are mainly based on a specific type of dynamic evolution of the agents' states known as *consensus schemes*. Each agent evolves its state by some sort of averaging of the states of its 'neighbours'. Each new state lies, in the convex hull of the previous averaged ones so that the limit value is the same for all the agents, under certain communication criteria [1].

1.1. Delayed systems and related literature. Time delays are inevitable in the study of real-world networked systems. These are the result of either finite speed of information propagation between agents, known as *communication delays*, or finite speed of information processing, known as *input delay*. In both cases, delays tend to weaken the performance of the system or in some cases destabilize it ([15]).

Furthermore it is also very important to study not only the problem of asymptotic behavior of such distributed systems but also the rate at which this behavior is revealed. It is very desirable to obtain mathematically tractable expressions on the

Key words and phrases. Consensus Dynamics, Fixed Point Theory, Contraction Mappings, Delays.

¹csomarak@umd.edu

²baras@umd.edu

³evripar@umd.edu

effect of delays in the performance of these dynamical systems as a function of the rest of the system's parameters.

Nevertheless, there are no strong results in the literature concerning the effect of delays in distributed consensus systems. To the best of our knowledge we mention a number of relative results. A simple delayed consensus algorithm was proposed and discussed in [15] where the model

$$\dot{x}_i(t) = \sum_j a_{ij}(x_j(t - \tau) - x_i(t - \tau))$$

With $\tau > 0$ constant and uniform for all agents, a frequency method analysis was carried through. The authors used frequency methods to show that a necessary and sufficient condition for convergence is $\tau < \frac{2\pi}{\lambda_N}$ where λ_N is the largest eigenvalue of the Laplacian matrix (see section A of the Appendix). The problem with this method is that it does not apply if the delays are multiple and incommensurate, or if the system is time varying or even non-linear. In [27, 1] the authors consider a discrete time version of Eq. (1) without processing delays and with time-varying information propagation delays $\tau(t)$. On condition that the delay is uniformly upper bounded, the strategy of attacking this problem is to extend the state space by adding artificial agents which played no actual role in the dynamics other than transmitting a pre-described delayed version of an agent's state. This approach, it is unclear how it would work in a continuous time system, unless the latter one is discretized. Next, in [16] the authors discuss the convergence properties of a non-linear model which has the form

$$\dot{x}_i = \sum_j a_{ij} f_{ij}(x_j(t - \tau) - x_i)$$

Using passivity assumptions on f_{ij} they apply invariance principles to derive delay-independent convergence results. The main setback of that approach, however, is that nothing can be said for either the rate of convergence to the consensus point or the consensus point itself. Another similar argument is made in [11] for a the linear consensus model with information propagation delays. The author is based on the linearity argument to conclude exponential stability from asymptotic stability.

The last family of models concerns rendezvous type of algorithms. In [14] the authors propose a second order consensus based algorithms, where agents asymptotically meet in a common place as their speed vanishes to zero. This algorithm is of the form

$$\dot{v}_i(t) = -cv_i(t) + \sum_j a_{ij}(r_j(t - \tau_j^i) - r_i(t))$$

The authors establish a Lyapunov-Krasovskii argument based on Invariance Principles and give delay-dependent results. Again, nothing can be said about the rate of convergence of this system since Invariance principles guarantee convergence.

1.2. Contribution. We contribute to the problem of delayed consensus networks by introducing FPT techniques to the stability analysis of a fully delayed linear time invariant consensus network, i.e. a network where each agent averages the delayed information of its neighbors with the delayed version of its own information. This approach stands in contrast to both frequency and Lyapunov methods used so far in the literature [11, 16, 15].

We will exploit the solution kernels of the un-delayed algorithm Eq. (17) to construct a solution operator which will proven to be a contraction in an especially designed complete metric space, each member of which converges to a constant value with prescribed rate of convergence. This rate can be explicitly stated as a function of the system's parameters. Then by the Contraction Mapping Principle

(see Theorem B of the Appendix) the solution operator has a unique fixed point. This is a de-facto solution of the stability problem.

Through examples we will outline our method and at the same time we will argue that either a Lyapunov or a Frequency approach for the stability of the proposed model is too restricting either on the assumptions for the communication graph or on the assumptions of the delays and in any case sheds little light upon the critical quantities associated with the asymptotic behavior of the solution, except for very special cases (of particular symmetry or increased connectivity [13, 24]).

This work is an outgrowth of [26] where a simpler model was discussed in the same framework. Here we study a more general model, a special case of which is the aforementioned work. Here we will derive convergence results under much milder assumptions than the ones imposed in [26] using a more elaborate form of the solution representation and a different metric.

1.3. Lyapunov and Fixed Points. Comparing Lyapunov and FPT is of interest in applied problems, only. It is very well known that a strong type of stability implies the existence of a Lyapunov function. However it is the discovery of this function the actual and in many cases of utmost difficulty problem. But once a function has been found, there is an abundance of theorems and techniques to extract fruitful information about properties of the solutions of the system.

In a similar vein, the problem in FPT is to find a suitable expression of the solution that will act as a map of one functional space to another. A typical way is to invert the differential operator using a variation of constants formula. Then an appropriately selected topological space and a metric function need to be considered. In our work this pair must constitute a complete metric space. Then the last two steps is to show that the defined operator is a function that maps the metric space into itself and it is a contraction under the defined metric. Then the Contraction Mapping Principle guarantees the existence of a unique fixed point.

It can be easily deduced that the real problem is to come up with a clever representation of the solution. Once such an expression is found there are numerous of fixed point theorems and functional spaces into which one may prove the existence of a fixed point. Then is the properties of the functional space in which the solution has been found that reflect the properties of the solution; in our case it is the stability and the rate of convergence. The interested reader is referred to the seminal monograph of T.A. Burton [3].

1.3.1. Fixed Point Approach in consensus systems. Distributed consensus algorithms attain a number of interesting properties which motivate this continuous interest for their study. At first they do not meet the conventional definition of asymptotic stability as any constant value $x_i \equiv k$ for all i is a solution. Investigators rather talk about asymptotically constant solutions of such systems. Next, it is the distributed nature of these dynamics that makes it difficult for the investigator to establish an estimate of use on the rate of convergence to the consensus state. In the case of linear time invariant dynamics, Algebraic Graph Theory methods gives a perfect solution on the problem when no delays are considered. Then one can consider the delayed system as a perturbation of the un-delayed one. This approach however puts massive restrictions on the allowed delayed bounds (see [26]). Another way to represent a solution is by a mere inversion of the differential operator with the use of the Variation of Constants formula for functional differential equations. This representation however captures local dynamical behavior only and cannot be used

alone. It is the combination of these two representations that can be used as a non-trivial form of solution. Let us illustrate our approach through a brief study of the following scalar problem

Example 1.1. Fix $a, \tau > 0$ and consider the initial value problem

$$\begin{cases} \dot{x}(t) &= -ax(t) + ax(t - \tau), t \geq 0 \\ x(t) &= \phi(t), t \in [-\tau, 0] \end{cases}$$

where $x \in \mathbb{R}$ and without loss of generality we set $\phi(t_0) + a \int_{-\tau}^0 \phi(s) ds = 0$. Then two non-trivial forms of the solution are

$$\begin{aligned} x(t) &= -a \int_{t-\tau}^t x(s) ds \\ x(t) &= e^{-a(t-t_0)} x(t_0) + \int_{t_0}^t e^{-a(t-s)} ax(s - \tau) ds, \forall t \geq t_0 \geq 0 \end{aligned}$$

It can be shown that the first expression defines a contraction (see Appendix B) in a complete metric space of functions that converge to a fixed constant as $t \rightarrow \infty$ regardless of the sign of a , if $|a\tau| < 1$ [3]. The second expression exploits the diffusive nature (i.e. $a > 0$) but it is not a contraction in any metric space of interest, exactly because the kernel e^{-at} is not strong enough to control the magnitude of the "perturbation" $ax(t - \tau)$. A combination of the two solutions, however, yields a contraction operator without the need to suppress τ , whenever $a > 0$. Indeed, from the second expression take $t - t_0 = \tau$ it is not hard to see that

$$x(t) = e^{-a\tau} x(t - \tau) + \int_{t-\tau}^t e^{-a(t-s)} ax(s - \tau) ds = \int_{t-2\tau}^{t-\tau} a(e^{-a(t-s-\tau)} - e^{-a\tau}) x(s) ds$$

This form of the solution defines a contractive operator for any $\tau \in [0, \infty)$. Then introducing a weighted function space we can establish an estimate on the rate of convergence that depends explicitly on a and τ .

In the multi-dimensional case a number of crucial, but straightforward, modifications need to be made. Unfortunately, multiple dimensions invoke drawbacks in the strength of the obtained results.

1.4. Organization of the paper. In Section 2, we introduce the main notations and definitions that we will use throughout the paper. In Section 3, we define our problem; pose the assumptions and state our main result. In section 4, we prove our main result. The proof is an application of the Contraction Mapping Principle and it is separated in several small steps. In Section 5, we illustrate our results with a number of examples and simulations. A thorough discussion on our result is held in Section 6. We analyse the effect of the assumptions, the overall advantages and disadvantages of this approach and consequently prospects for future work along these lines. Furthermore we propose a number of interesting variations of our model where such techniques can be adapted together with the necessary modifications. Fundamentals of Algebraic Graph Theory and auxiliary results on linear consensus dynamics are provided in Appendix A. Elements of FPT are presented in Appendices B. Auxilliary results from the theory of linear inequalities are presented in Appendix C. All the proofs are put in Appendix D.

2. Notations and Definitions. By $N < \infty$ we denote the number of agents whereas the set of agents is denoted by $[N] := \{1, \dots, N\}$. Each agent $i \in [N]$ is associated with a real quantity $x_i \in \mathbb{R}$. The Euclidean vector space \mathbb{R}^N is the state

space of the system with the state vectors $\mathbf{x} = (x_1, \dots, x_N)^T$. For $\mathbf{x} \in \mathbb{R}^N$ we define the *weighted* maximum norm $\|\mathbf{x}\|_{\mathbf{q}} := \max_i q_i |x_i|$ for some fixed $\mathbf{q} \in \mathbb{R}_+^N$, together with the corresponding induced matrix norm. By $\Re(z)$ we understand the real part of $z \in \mathbb{C}$ and by $\mathbb{1}$ we understand the N dimensional, column vector of all ones. The subspace of \mathbb{R}^N of interest is defined by $\Delta = \{\mathbf{y} \in \mathbb{R}^N : \mathbf{y} = \mathbb{1}k, k \in \mathbb{R}\}$ and it is called the consensus space. Following the standard framework of the theory of retarded functional differential equations, the symbols “ \cdot ” and $\frac{d}{dt}$ stand for the right hand-side derivative [6]. Also recall the big-oh notation where $r(y) = \mathcal{O}(s(y))$ implies $|r(y)| \leq B|s(y)|$ for some $B > 0$ and $y > y_0$. The Kronecker symbol is denoted by δ_{ij} . Finally, $C^0(I, \mathbb{R}^N)$ is the set of \mathbb{R}^N -valued continuous functions defined in I and L^1_I is the set of absolute integrable functions in I . For the rest of the notation the reader is referred to the Appendices A and B.

By *delay-dependent results* we mean that the result is valid if the maximum delay is bounded above by a number depending on the undelayed systems parameters. By *delay-independent results* we mean that the result is valid regardless of the value of the delay so long as the latter is finite.

3. The Model. Fix $N < \infty$ and $\tau_j^i, \sigma_i^j \in [0, \infty)$ with $\tau := \max_{i,j} \{\tau_j^i\}$, $\sigma := \max_{i,j} \{\sigma_i^j\}$ and consider the following initial value problem:

$$\begin{aligned} \dot{x}_i(t) &= \sum_{j \in N_i} a_{ij} (x_j(t - \tau_j^i) - x_i(t - \sigma_i^j)), \quad t \geq 0 \\ x_i(t) &= \phi_i(t), \quad t \in [-\max\{\sigma, \tau\}, 0] \end{aligned} \quad (1)$$

where N_i is the subset of $[N]$ for which $a_{ij} \geq 0$ and $\phi_i \in C^0([-\max\{\sigma, \tau\}, 0], \mathbb{R})$ are given initial functions.

3.0.1. *The Assumptions.* Recall the notations and definitions from Section 2 and from the Appendix A.

Assumption 3.1. *The associated communication graph, $G = ([N], E, W)$, is routed-out branching.*

Assumption 3.2. *The system parameters $a_{ij}, \tau_j^i, \sigma_i^j$ and \mathbf{c} satisfy*

$$1 + \sum_{i=1}^N c_i \sum_{j \in N_i} a_{ij} (\tau_j^i - \sigma_i^j) > 0$$

By Assumption 3.1 we conclude that $e^{-Lt}L = (e^{-Lt} - \mathbb{1}\mathbf{c}^T)L$ is a function matrix with elements that converge to zero exponentially fast with rate $\Re\{\lambda\}$ (see also Proposition A.1). Let κ_{il} be the value of the $(i, l)^{th}$ element of $e^{-Lt}L$. Define

$$g_{il}(\gamma) := \sup_{t \geq \tau} \int_0^{t-\tau} |\kappa_{il}(t - \tau - s)| e^{\gamma(t-s)} ds \quad (2)$$

a quantity that is well-defined for $\gamma < \Re\{\lambda\}$ and

$$h_{l,j,i}(\gamma) := |a_{lj} - a_{ij}| \frac{e^{\gamma\tau_j^l} - 1}{\gamma} + a_{ij} \frac{e^{\gamma \max\{\tau_j^i, \tau_j^l\}} - e^{\gamma \min\{\tau_j^i, \tau_j^l\}}}{\gamma} \quad (3)$$

Define

$$f_{ij}(\gamma) := a_{ij} e^{\gamma \tau_j^i} \frac{1 - e^{-(d_i - \gamma)\tau}}{d_i - \gamma} - a_{ij} e^{-(d_i - \gamma)\tau} \frac{e^{\gamma \tau_j^i} - 1}{\gamma} + e^{-d_i \tau} \sum_{l \neq i} h_{l,j,i}(\gamma) g_{il}(\gamma) + e^{-d_i \tau} \sum_l g_{ij}(\gamma) a_{jl} \frac{e^{\gamma \sigma_j^l} - 1}{\gamma}$$

if $i \neq j$ and

$$-f_{ii}(\gamma) := 1 - \sum_{l \neq i} a_{il} \frac{e^{\gamma \sigma_i^l} - 1}{\gamma} \left(1 + d_i \frac{1 - e^{-(d_i - \gamma)\tau}}{d_i - \gamma} \right) - e^{-d_i \tau} \sum_{l \neq i} h_{l,i,i}(\gamma) g_{il}(\gamma) - e^{-d_i \tau} \sum_{l \neq i} g_{ii}(\gamma) a_{il} \frac{e^{\gamma \sigma_i^l} - 1}{\gamma}$$

otherwise and consider the matrix $\mathbf{F}(\gamma) = [f_{ij}(\gamma)]$.

Assumption 3.3. *There exists $\mathbf{q} \in \mathbb{R}_+$ such that for all $i \in [N]$ so that*

$$\mathbf{F}(0) \frac{1}{\mathbf{q}} < 0$$

where $\frac{1}{\mathbf{q}} := (1/q_1, \dots, 1/q_N)$.

We state now the main result of this paper.

Theorem 3.4. *Under Assumptions 3.1, 3.2 and 3.3, all states in Eq. (1) converge to*

$$k = \frac{\sum_{i=1}^N c_i [\phi_i(0) + \sum_{j \in N_i} a_{ij} (\int_{-\tau_j^i}^0 \phi_j(s) ds - \int_{-\sigma_i^j}^0 \phi_i(s) ds)]}{1 + \sum_{i=1}^N c_i \sum_{j \in N_i} a_{ij} (\tau_j^i - \sigma_i^j)} \quad (4)$$

exponentially fast with rate $\gamma \in (0, \min\{d_i, \Re\{\lambda\}\})$ which satisfies

$$\mathbf{F}(\gamma) \frac{1}{\mathbf{q}} \leq 0.$$

Assumption 3.1 is a minimal connectivity condition which requires that there can be at most one $i^* \in [N]$ with $d_{i^*} = 0$, usually called the leader of the group, in which case for $\mathbf{c} = (c_1, \dots, c_N)^T$ the the left eigenvector of $L = L_G$ associated with $\lambda_1 = 0$, we have $\mathbf{c} = e_{i^*}$. This means that i^* remains unaffected and all other agents' states asymptotically converge to the state of i^* . For the rest of the paper we will quietly assume that $d_i > 0, \forall i \in [N]$ without essentially effecting any of the derived results. The assumptions on the boundedness of the delays and the convergence of the disturbances are needed in order to maintain exponential convergence of the system to the consensus state. This state is, in turn, analytically calculated from the initial conditions and the topological structure of the graph. Imploring FPT methods in the study of the stability of solutions of the initial value problem (1), we extract sufficient conditions for their asymptotic behavior of it's solutions. Although our analysis is in principle based on the smallness of the maximum allowed delays σ and τ (Assumption 3.3) in the Examples section we will study cases where the stability results become delay-independent.

4. Proof of the main result. The proof is an application of the Contraction Mapping Principle (Theorem B.2 of the Appendix B) and it is separated into several propositions every proof of which is put in the Appendix D. Most of the quantities to be mentioned are defined in the Appendices and are taken from there freely.

4.1. Preparation of the solution operator. The solutions of the system (1) can be expressed into two different ways. We will exploit both of them to create a new one which will be our main operator. For $i \in [N]$ we write from Eq. (1):

$$\dot{x}_i = \sum_{j \in N_i} a_{ij}(x_j - x_i) - \sum_{j \in N_i} a_{ij} \frac{d}{dt} \left[\int_{t-\tau_j^i}^t x_j(s) ds - \int_{t-\sigma_i^j}^t x_i(s) ds \right]$$

In vector form this equation reads

$$\dot{\mathbf{x}} = -L\mathbf{x} - \sum_{i,j} A_{ij} \frac{d}{dt} \int_{t-\tau_j^i}^t \mathbf{x}(s) ds + \sum_{i,j} B_{ij} \frac{d}{dt} \int_{t-\sigma_i^j}^t \mathbf{x}(s) ds$$

where $A_{ij} = [a_{lk} \delta_{li} \delta_{kj}]$ and $B_{ij} = [a_{lk} \delta_{li} \delta_{kj}]$. Using the variation of constants and the integration by parts formulae we see that the solution of (1) satisfies

$$\begin{aligned} \mathbf{x}(t) &= e^{-Lt} \phi(0) - \int_0^t e^{-L(t-s)} \sum_{i,j} \left[A_{ij} \frac{d}{ds} \int_{s-\tau_j^i}^s \mathbf{x}(w) dw - B_{ij} \frac{d}{ds} \int_{s-\sigma_i^j}^s \mathbf{x}(w) dw \right] ds \\ &= e^{-Lt} \mathbf{r}_0 - \sum_{i,j} \left[A_{ij} \int_{t-\tau_j^i}^t \mathbf{x}(s) ds - B_{ij} \int_{t-\sigma_i^j}^t \mathbf{x}(s) ds \right] + \\ &+ \int_0^t e^{-L(t-s)} L \sum_{i,j} \left[A_{ij} \int_{s-\tau_j^i}^s \mathbf{x}(w) dw - B_{ij} \int_{s-\sigma_i^j}^s \mathbf{x}(w) dw \right] ds \end{aligned} \quad (5)$$

where

$$\mathbf{r}_0 := \phi(0) + \sum_{i,j} \left[A_{ij} \int_{-\tau_j^i}^0 \phi(s) ds - B_{ij} \int_{-\sigma_i^j}^0 \phi(s) ds \right] \quad (6)$$

An alternative way, to express the solution of (1) in vector form is

$$\dot{\mathbf{x}}(t) = -D\mathbf{x}(t) + \sum_{i,j} A_{ij} \mathbf{x}(t - \tau_j^i) + \sum_{i,j} B_{ij} \frac{d}{dt} \int_{t-\sigma_i^j}^t \mathbf{x}(s) ds$$

and inversion from $t - \tau$ to t yields

$$\mathbf{x}(t) = e^{-D\tau} \mathbf{x}(t - \tau) + \int_{t-\tau}^t e^{-D(t-s)} \sum_{i,j} \left[A_{ij} \mathbf{x}(s - \tau_j^i) + B_{ij} \frac{d}{ds} \int_{s-\sigma_i^j}^s \mathbf{x}(w) dw \right] ds \quad (7)$$

It is our intention to combine these two forms of solution to a new one. From (7)

$$\begin{aligned} \mathbf{x}(t) &= e^{-D\tau} \mathbf{x}(t - \tau) + \int_{t-\tau}^t e^{-D(t-s)} \sum_{i,j} \left[A_{ij} \mathbf{x}(s - \tau_j^i) + B_{ij} \frac{d}{ds} \int_{s-\sigma_i^j}^s \mathbf{x}(w) dw \right] ds \\ &= e^{-D\tau} \mathbf{x}_{(5)}(t - \tau) + \int_{t-\tau}^t e^{-D(t-s)} \sum_{i,j} \left[A_{ij} \mathbf{x}(s - \tau_j^i) + B_{ij} \frac{d}{ds} \int_{s-\sigma_i^j}^s \mathbf{x}(w) dw \right] ds \end{aligned}$$

So that for $t \geq \tau$

$$\begin{aligned} \mathbf{x}(t) &:= e^{-D\tau} e^{-L(t-\tau)} \mathbf{r}_0 + \\ &+ \sum_{i,j} \left[\int_{t-\tau-\tau_j^i}^{t-\tau} \left(e^{-D(t-\tau_j^i-s)} - e^{-D\tau} \right) A_{ij} \mathbf{x}(s) ds + \int_{t-\tau}^{t-\tau_j^i} e^{-D(t-\tau_j^i-s)} A_{ij} \mathbf{x}(s) ds \right] + \\ &+ \sum_{i,j} \left[B_{ij} \int_{t-\sigma_i^j}^t \mathbf{x}(w) dw ds - \int_{t-\tau}^t e^{-D(t-s)} D B_{ij} \int_{s-\sigma_i^j}^s \mathbf{x}(w) dw ds \right] + \\ &+ e^{-D\tau} \int_0^{t-\tau} e^{-L(t-\tau-s)} L \sum_{i,j} \left[A_{ij} \int_{s-\tau_j^i}^s \mathbf{x}(w) dw - B_{ij} \int_{s-\sigma_i^j}^s \mathbf{x}(w) dw \right] ds \end{aligned} \quad (8)$$

4.2. The space of solutions and the solution operator. The linearity of the problem allows us to work in complete metric spaces and apply Banach's Contraction Principle [B.2](#). For asymptotic stability with prescribed convergence estimate, though, we will need an especially designed metric space.

4.2.1. The space of solutions. Fix $\psi \in C^0([-p, \tau], \mathbb{R}^N)$, $\gamma > 0$, $k \in \mathbb{R}$. By $\mathbb{B} = C^0([-p, \infty), \mathbb{R}^N)$ we define the set of continuous bounded (in the sense of the supremum norm) functions defined in $[-p, \infty)$ and which take values in \mathbb{R}^N . Define

$$\mathbb{M} = \left\{ \mathbf{y} \in \mathbb{B} : \mathbf{y} = \psi|_{[-p, \tau]}, \sup_{t \geq \tau} e^{\gamma t} \|\mathbf{y}(t) - \mathbb{1}k\|_{\mathbf{q}} < \infty \right\} \quad (9)$$

together with the function

$$\rho(\mathbf{y}_1, \mathbf{y}_2) := \sup_{t \geq \tau} e^{\gamma t} \|\mathbf{y}_1(t) - \mathbf{y}_2(t)\|_{\mathbf{q}} \quad (10)$$

This is the space of \mathbb{R}^N -valued continuous functions each member of which agrees on $[-p, \tau]$ with a prescribed function ϕ and it converges to $\mathbb{1}k$ exponentially fast with rate γ . We can readily attach this space to our problem by picking $\tau = \max_{i,j} \tau_j^i$, $p = \max\{\sigma, \tau\}$, $\psi(t) = \phi(t)$ for $t \in [-p, 0]$ and $\psi(t) = \mathbf{x}_{(1)}(t)$ $t \in [0, \tau]$, i.e. the unique solution of [\(1\)](#) in $[0, \tau]$. This formulation considers the existence and uniqueness for the solution of [\(1\)](#) in $[0, \tau]$ which is hardly an assumption due to the linearity of the model. The next result is of fundamental importance.

Proposition 4.1. (\mathbb{M}, ρ) constitutes a complete metric space.

4.2.2. The solution operator. Next we define $\mathcal{P} : \mathbb{M} \rightarrow \mathbb{B}$ as follows:

$$(\mathcal{P}\mathbf{x})(t) = \begin{cases} \psi(t), & t \in [-\max\{\tau, \sigma\}, \tau] \\ \mathbf{x}_{(8)}(t), & t \geq \tau. \end{cases} \quad (11)$$

where $\mathbf{x}_{(8)}(t)$ is the right hand-side of Eq. [\(8\)](#).

4.3. Stability Analysis. The next step is to prove that \mathcal{P} maps \mathbb{M} into itself. The next lemma shows that this is true for specific k and γ .

Proposition 4.2. $\mathcal{P} : \mathbb{M} \rightarrow \mathbb{M}$ if k is defined as in [\(4\)](#) and $\gamma < \Re(\lambda)$.

The number k in [\(4\)](#) is the consensus point which is, as expected, a function only of the parameters of the system, the initial data and it is well-defined by Assumption [3.2](#).

4.3.1. The Contraction Property of \mathcal{P} . The final step is to show that \mathcal{P} satisfies the contraction property for some γ .

Lemma 4.3. Under Assumption [3.3](#), $\mathcal{P} : \mathbb{M} \rightarrow \mathbb{M}$ is a contraction for some $\gamma \in (0, \min\{d_i, \Re(\lambda)\})$.

We summarize the above results to conclude:

Proof of Theorem [3.4](#). This is an immediate result of Propositions [4.1](#), [4.2](#), Lemma [4.3](#) and Theorem [B.2](#). \square

5. Examples and Simulations.

Example 5.1. For $N = 2$, Eq. (1) becomes

$$\begin{aligned} \dot{x}(t) &= -ax(t - \sigma_1) + ay(t - \tau_1) \\ \dot{y}(t) &= -by(t - \sigma_2) + bx(t - \tau_2) \end{aligned} \quad (\text{Ex.1})$$

where with a, b so that Assumption 3.1 holds and $\sigma_i, \tau_i \geq 0$. Without loss of generality we take $\tau_1 \geq \tau_2$. Set also for convenience $\Lambda = \frac{ab}{a+b}$. Since $\mathbf{c} = \{\frac{b}{a+b}, \frac{a}{a+b}\}$ Assumption 3.2 requires

$$1 + \Lambda(\tau_1 + \tau_2 - \sigma_1 - \sigma_2) > 0 \quad (12)$$

Next,

$$e^{-Lt}L = \begin{bmatrix} a & -a \\ -b & b \end{bmatrix} e^{-(a+b)t}$$

The matrix $G(\gamma) = \{g_{ij}(\gamma)\}$ reads

$$G(\gamma) = \begin{bmatrix} a & a \\ b & b \end{bmatrix} \frac{e^{\gamma\tau_1}}{a+b-\gamma}$$

whenever $\gamma < a + b$. Then we consider the elements of $\mathbf{F}(\gamma)$

$$\begin{aligned} f_{12} &:= ae^{-\gamma\tau_1} \frac{1 - e^{-(a-\gamma)\tau_1}}{a-\gamma} - ae^{-(a-\gamma)\tau_1} \frac{e^{\gamma\tau_1} - 1}{\gamma} + ae^{-a\tau_1} \frac{e^{\gamma\tau_1} - 1}{\gamma} \frac{a}{a+b-\gamma} e^{\gamma\tau_1} + e^{-a\tau_1} \frac{abe^{\gamma\tau_1}}{a+b-\gamma} \frac{e^{\gamma\sigma_2} - 1}{\gamma} \\ -f_{11} &:= \left(1 - a \frac{e^{\gamma\sigma_1} - 1}{\gamma} \left(1 + \frac{a}{a-\gamma} (1 - e^{-(a-\gamma)\tau_1}) \right) - e^{-a\tau_1} \Lambda \left(\frac{e^{\gamma\tau_2} - 1}{\gamma} + \frac{a}{b} \frac{e^{\gamma\sigma_1} - 1}{\gamma} \right) \right) \\ f_{21} &:= be^{-\gamma\tau_2} \frac{1 - e^{-(b-\gamma)\tau_1}}{b-\gamma} - be^{-(b-\gamma)\tau_1} \frac{e^{\gamma\tau_2} - 1}{\gamma} + be^{-b\tau_1} \frac{e^{\gamma\tau_2} - 1}{\gamma} \frac{b}{a+b-\gamma} e^{\gamma\tau_1} + e^{-b\tau_1} \frac{abe^{\gamma\tau_1}}{a+b-\gamma} \frac{e^{\gamma\sigma_1} - 1}{\gamma} \\ -f_{22} &:= \left(1 - b \frac{e^{\gamma\sigma_2} - 1}{\gamma} \left(1 + \frac{b}{b-\gamma} (1 - e^{-(b-\gamma)\tau_1}) \right) - e^{-b\tau_1} \Lambda \left(\frac{e^{\gamma\tau_1} - 1}{\gamma} + \frac{b}{a} \frac{e^{\gamma\sigma_2} - 1}{\gamma} \right) \right) \end{aligned}$$

and for $\gamma \downarrow 0$, the elements of $\mathbf{F}(0)$

$$\begin{aligned} f_{12} &:= (1 - e^{-a\tau_1} - \Lambda\tau_1 e^{-a\tau_1} + e^{-a\tau_1} \Lambda\sigma_2) \\ -f_{11} &:= \left(1 - a\sigma_1(2 - e^{-a\tau_1}) - e^{-a\tau_1} \Lambda \left(\tau_2 + \frac{a}{b} \sigma_1 \right) \right) \\ f_{21} &:= (1 - e^{-b\tau_1} - \Lambda\tau_2 e^{-b\tau_1} + e^{-b\tau_1} \Lambda\sigma_1) \\ -f_{22} &:= \left(1 - b\sigma_2(2 - e^{-b\tau_1}) - e^{-b\tau_1} \Lambda \left(\tau_1 + \frac{b}{a} \sigma_2 \right) \right) \end{aligned}$$

Then Assumption 3.3 requires to find $q_1, q_2 > 0$ so that

$$f_{12} \frac{1}{q_2} < (-f_{11}) \frac{1}{q_1} \quad \text{and} \quad f_{21} \frac{1}{q_1} < (-f_{22}) \frac{1}{q_2}$$

a set of linear inequalities which is consistent if and only if

$$f_{11}f_{22} > f_{12}f_{21} \quad (13)$$

whenever $f_{ii} < 0, f_{ij} > 0$. Eqs. (12) and (13) describe the allowed bounds for the processing and propagation delays.

As a numerical application we take $a = 0.5$ and $b = 1.3$, $\tau_1 = 1$ $\tau_2 = 0.2$ and $\sigma_1 = \sigma_2 = 0.215$. Then Eq. (13) remains consistent for $\gamma \leq 0.045$ which is our estimate for the rate of convergence. Next we focus on two extreme cases:

No propagation delays. In this case the bounds are rightfully much stricter. As the necessary condition by Assumption 3.2 suggests it is impossible to obtain stability for any bounded σ . The conditions imposed are

$$\sigma_1 + \sigma_2 < \frac{1}{\Lambda}$$

by Assumption 3.2 and

$$1 + 2ab\sigma_1\sigma_2 > \frac{2(a^2\sigma_1 + b^2\sigma_2) + ab(\sigma_1 + \sigma_2)}{a + b}$$

from condition (13). More specifically, for $\sigma_1 = \sigma_2 = \sigma$ we ask

$$\sigma < \frac{1}{2} \frac{a+b}{ab}, \quad \text{and} \quad \sigma^2 - \frac{a^2 + b^2 + (a+b)^2}{2ab(a+b)}\sigma + \frac{1}{2ab} > 0.$$

No processing delays. If we ignore the processing delays, the requirement of Assumption 3.2 is automatically satisfied while from condition (13) it is simplified to

$$(1 - \Lambda\tau_1 e^{-b\tau_1})(1 - \Lambda\tau_2 e^{-a\tau_1}) > (1 - e^{-a\tau_1} - \Lambda\tau_1 e^{-a\tau_1})(1 - e^{-b\tau_1} - \Lambda\tau_2 e^{-b\tau_1})$$

Consequently, so long as the above inequality is satisfied, there are always $q_1, q_2 > 0$ and hence a weighted metric ρ so that the operator as defined in (11) is a contraction in (\mathbb{M}, ρ) .

Note that the above inequality does not hold for any a, b, τ_1, τ_2 . It is true however either when $a = b > 0$ and arbitrary $\tau_1, \tau_2 \geq 0$ or when $\tau_1 = \tau_2$ and arbitrary $a, b \geq 0$. This allows us to establish bounds around any nominal value w^* in the vicinity of which for any given τ_1, τ_2 there exists a radius r so that a, b can lie in $B(w^*, r)$ and similarly for τ_1, τ_2 . In our numerical example we take $\sigma_1 = \sigma_2 = 0$ and the rate estimate we get is $\gamma = 0.4545$.

Remark 5.2. In [26] we considered Eq. 1 with $\sigma = 0$ and we used an operator based exclusively on the solution expression of Eq. (5). Then under the L^1 norm, $\|\cdot\|_1$ we derived the following contraction condition

$$\tilde{A}\tau \left(1 + \frac{\sqrt{N}\|L\|_1}{\lambda_2}\right) < 1 \tag{14}$$

where $\tilde{A} = \sum_{i=1}^N \sum_{j \in N_i} a_{ij}$ and $a_{ij} = a_{ji}$. It should be intuitively clear that Eq. 14 is much stricter than the condition (3.3) and we will illustrate this difference within this example. Applying the bound (14) stability is ensured if

$$\tau < \frac{1}{2(1 + \sqrt{2})a}$$

Remark 5.3. Taking $a = b$ and $\tau_i = \sigma_i$ we can compare our results with [15] where the necessary and sufficient condition is

$$\tau < \frac{2\pi}{\lambda_N} = \frac{\pi}{a}$$

while our bounds ask

$$\tau < \frac{1}{2a}$$

Remark 5.4. We conclude this example by mentioning a candidate Lyapunov functional in the case $\tau_1 = \tau_2 = \tau$.

$$V(x, y) = bx^2 + ay^2 + ab \int_{t-\tau}^t x^2(s) + y^2(s) ds$$

This is a functional on \mathbb{R}^2 it is obviously continuous and

$$\dot{V} = -ab[(x - y_t)^2 + (y - x_t)^2]$$

Then the set, S , such that $\dot{V}(t) \equiv 0$ is the one where $x(t) = y(t - \tau)$ and $y(t) = x(t - \tau)$ for all t . The largest subset of S that is invariant with respect to these dynamics is Δ . Then standard Invariance Theory arguments yield asymptotic convergence to the consensus subspace (Section 5.3 of [6]) independent of the magnitude of the delay (but without any estimate on the rate of convergence).

Example 5.5 (A Delayed Complete Graph.). Let (1) with $a_{ij} \equiv \theta$, $\sigma_i^j \equiv 0$ $\tau_j^i = \tau_i$. Under this communication scheme every agent is connected with everyone else with identical connectivity weight and they receive the signals with propagation delay that depends on agent i only. The complete graph on N agents has the edge set $E = \{(i, j) : i \neq j\}$. The Laplacian of this graph has the spectrum $\lambda_1 = 0, \lambda_i = N\theta \mathbb{1}_{i=2}^N$. Then

$$[e^{-Lt}L]_{ij} = \begin{cases} \theta \frac{N-1}{N} e^{-N\theta t}, & i = j \\ \theta \frac{-1}{N} e^{-N\theta t}, & i \neq j \end{cases}$$

and $d_i \equiv (N-1)\theta$ and the left eigenvector of L is $\mathbf{c} = \mathbb{1} \frac{1}{N}$. Assumption 3.2 is automatically satisfied whereas for Assumption 3.3 we have

$$f_{ij}(0) = \frac{1}{N-1} (1 - e^{-(N-1)\theta\tau}) - \theta\tau_i e^{-(N-1)\theta\tau} + \frac{\theta}{N^2} e^{-(N-1)\theta\tau} \sum_{l \neq i} (\max\{\tau_i, \tau_l\} - \min\{\tau_i, \tau_l\})$$

$$f_{ii}(0) = -1 + \frac{\theta}{N^2} e^{-(N-1)\theta\tau} \sum_{l \neq i} \tau_l$$

We will apply Theorem (C.1) and for this we will need the following condition. Let for $i \in [N]$, N_i denote the number of $l \neq i$ such that $\tau_l \geq \tau_i$. We ask that

$$1 + \frac{N^2 + 2N_i - N}{N} \theta\tau_i - \frac{N-1}{N^2} \theta \left[\sum_{l: \tau_l \geq \tau_i} \tau_l - \sum_{l: \tau_l < \tau_i} \tau_l \right] > 0 \quad (15)$$

Note that this condition is satisfied if for example $\tau_i \equiv \tau$. From Theorem (C.1) we choose $m = 2N$, $\mathbf{a}_i := (f_{i1}, \dots, f_{iN})$ for $i = 1, \dots, N$, $\mathbf{a}_i < 0$ elementwise for $i = N+1, \dots, 2N$, $\alpha_i = 0$ for $i = 1, \dots, N$ and $\alpha_i < 0$ for $i = N+1, \dots, 2N$. Then for the sake of contradiction if the second case holds there exist $\xi_i|_{i=1}^{2N} \geq 0$ so that at least one of $\xi|_{i \leq N}$ is positive and $\sum_{i=1}^{2N} \alpha_i \xi_i = \sum_{i=N+1}^{2N} \alpha_i \xi_i \leq 0$ and $\xi_i|_{i \geq N+1} \geq 0$. From the second condition

$$\sum_{i=1}^{2N} \mathbf{a}_i \xi_i = 0 \Rightarrow \sum_{j=1}^N f_{ij} \xi_j + \sum_{j=1}^N a_{ij} \xi_j = 0, \quad \forall i$$

Since the second part of the last equation is non-positive from the imposed condition (15) it can be verified that it implies $\sum_j f_{ij} < 0$ for all i , hence a contradiction because not all $\xi|_{i \leq N}$ can be zero. So there exists a set of positive numbers (i.e. a weighted metric) to satisfy the Assumption 3.3.

Example 5.6 (Uniform Delays in a topological Star Graph). We consider the star graph among N agents and we enumerate the central node of the graph to be the first agent. We take the communication weights identically equal to the unity. It

can be shown that

$$[e^{-Lt}L]_{ij} = \begin{cases} (N-1)e^{-Nt}, & i = j = 1 \\ -e^{-Nt}, & i = 1, j \neq 1 \text{ or } j = 1, i \neq 1 \\ \frac{1}{N-1}e^{-Nt} + \frac{N-2}{N-1}e^{-t}, & i = j, i \neq 1 \\ \frac{1}{N-1}e^{-Nt} - \frac{1}{N-1}e^{-t}, & o.w. \end{cases}$$

so that $g_{ii}(0) = \frac{N-1}{N}$ and $g_{ij}(0) = \frac{1}{N}$ otherwise. Assumption 3.2 asks for

$$1 + \frac{1}{N} \sum_{j \neq 1} (\tau_j^1 - \sigma_1^j) + \frac{1}{N} \sum_{j \neq 1} (\tau_1^j - \sigma_j^1) > 0, \quad i \in [N]$$

In particular if $\tau_j^i \equiv \tau$ and $\sigma_j^i \equiv 0$ the condition is automatically satisfied. Finally, for Assumption 3.3 we calculate the elements of $\mathbf{F}(0)$

$$f_{1j} = \begin{cases} -1 + \frac{N-1}{N}\tau e^{-(N-1)\tau} & j = 1 \\ \frac{1}{N-1}(1 - e^{-(N-1)\tau}) - \tau e^{-(N-1)\tau} + \frac{N-1}{N}\tau e^{-(N-1)\tau} & o.w. \end{cases}$$

and for $i > 1$

$$f_{ij} = \begin{cases} -1 & j = i \\ 1 - e^{-\tau} - \tau e^{-\tau} + \frac{\tau}{N}e^{-\tau} & j = 1 \\ \frac{\tau}{N}e^{-\tau} & o.w. \end{cases}$$

A simple calculation shows that $\sum_j f_{ij} < 0$ and the argument proceeds as in the Example 5.5.

Let us now turn to a numerical example where FPT methods suffer from the asymmetry of the delays and the weighted topology and hence the derived conditions ask for a very small bound on the maximum allowed delay τ .

Example 5.7. Consider a weighted graph with 4 nodes and the symmetric Laplacian matrix L

$$L = \begin{bmatrix} 6.3 & 0 & -2 & -4.3 \\ 0 & 4.8 & -3 & -1.8 \\ -2 & -3 & 6.1 & -1.1 \\ -4.3 & -1.8 & -1.1 & 7.2 \end{bmatrix}$$

We also assume the distribution of delays:

$$T_1 = \begin{bmatrix} 0 & 0.6241 & 0.9880 & 0.7962 \\ 0.5211 & 0 & 0.0377 & 0.0987 \\ 0.2316 & 0.3955 & 0 & 0.2619 \\ 0.4889 & 0.3674 & 0.9133 & 0 \end{bmatrix} \tau$$

with the control parameter τ . We calculate the allowed bound for Eq. (14) $\Re(\lambda) = 4.534$, $\|L\|_1 = 14.400$, $\bar{A} = 24.400$ and we obtain $\tau_{(14)} < 0.0057$. Now we will Apply

Theorem 3.4. We execute the following calculations.

$$\begin{aligned}
e^{-Lt}L &= \begin{bmatrix} -20.66 & -11.80 & 8.07 & 24.39 \\ -11.80 & -2.51 & 4.84 & 9.47 \\ 8.07 & 4.84 & -3.15 & -9.76 \\ 24.39 & 9.47 & -9.76 & -24.10 \end{bmatrix} e^{-4.53t} + \\
&+ \begin{bmatrix} 0.06 & -0.37 & 0.61 & -0.3 \\ -0.37 & 1.82 & -3.01 & 1.56 \\ 0.61 & -3.01 & 5.00 & -2.60 \\ -0.30 & 1.56 & -2.60 & 1.34 \end{bmatrix} e^{-8.23t} + \\
&+ \begin{bmatrix} 26.88 & 12.17 & -10.67 & -28.38 \\ 12.17 & 5.50 & -4.84 & -12.83 \\ -10.67 & -4.84 & 4.25 & 11.26 \\ -28.38 & -12.83 & 11.26 & 29.95 \end{bmatrix} e^{-11.63t}
\end{aligned}$$

where we used Putzer's Algorithm [18] and MAPLE. Next for $\tau_{\max} = 0.988\tau$ the matrix $G(0)$ is approximated as

$$G(0) = \begin{bmatrix} 2.454 & 1.603 & 0.993 & 2.999 \\ 1.603 & 0.413 & 0.494 & 1.215 \\ 0.993 & 0.494 & 0.535 & 1.502 \\ 2.999 & 1.215 & 1.502 & 2.826 \end{bmatrix}$$

also and with the use of MAPLE we calculate $\mathbf{F}(0)$ as a function of τ :

$$\mathbf{F} := [\mathbf{F}_1 : \mathbf{F}_2 : \mathbf{F}_3 : \mathbf{F}_4]$$

where

$$\begin{aligned}
\mathbf{F}_1 &= \begin{bmatrix} 6.76\tau e^{-6.22\tau} - 1 \\ 2.78\tau e^{-4.74\tau} \\ 0.32(1 - e^{-6.02\tau}) + 3.25\tau e^{-6.02\tau} \\ 0.59(1 - e^{-7.11\tau}) + 9.55\tau e^{-7.11\tau} \end{bmatrix}, \mathbf{F}_2 = \begin{bmatrix} 3.16\tau e^{-6.22\tau} \\ 1.38\tau e^{-4.74\tau} - 1 \\ 0.49(1 - e^{-6.02\tau}) + 2.72\tau e^{-6.02\tau} \\ 0.25(1 - e^{-7.11\tau}) + 5.68\tau e^{-7.11\tau} \end{bmatrix} \\
\mathbf{F}_3 &= \begin{bmatrix} 0.31(1 - e^{-6.22\tau}) + 6.00\tau e^{-6.22\tau} \\ 0.62(1 - e^{-4.74\tau}) + 11.39\tau e^{-4.74\tau} \\ 3.52\tau e^{-6.02\tau} - 1 \\ 0.16(1 - e^{-7.11\tau}) + 4.67\tau e^{-7.11\tau} \end{bmatrix}, \mathbf{F}_4 = \begin{bmatrix} 0.69(1 - e^{-6.22\tau}) + 15.16\tau e^{-6.22\tau} \\ 0.38(1 - e^{-4.74\tau}) + 5.47\tau e^{-4.74\tau} \\ 0.19(1 - e^{-6.02\tau}) + 3.38\tau e^{-6.02\tau} \\ 10.91\tau e^{-7.11\tau} - 1 \end{bmatrix}
\end{aligned}$$

then

$$\sum_j f_{ij} = \begin{bmatrix} 31.08\tau e^{-6.22\tau} - e^{-6.22\tau} \\ 21.02\tau e^{-4.74\tau} - e^{-4.74\tau} \\ 9.35\tau e^{-6.02\tau} - e^{-6.02\tau} \\ 30.81\tau e^{-7.11\tau} - e^{-7.11\tau} \end{bmatrix}$$

and the first τ such that $\sum_i f_{ij} = 0$ is $\tau^* = 0.0414$, so stability is ensured for $\tau < \tau^*$ with the same argument as in the Example 5.5. This is an improvement of the bound in Eq. (14) by almost an order of 10.

Example 5.8. [A simulation example] Consider the weighted network

$$L = \begin{bmatrix} 0.2 & -0.2 & 0 \\ -0.1 & 0.1 & 0 \\ -0.4 & 0 & 0.4 \end{bmatrix}$$

and the distribution of the processing and propagation delays

$$\Sigma(\sigma) = \begin{bmatrix} 0 & \frac{\sigma}{2} & 0 \\ \frac{\sigma}{6} & 0 & 0 \\ \frac{\sigma}{10} & \frac{\sigma}{7} & 0 \end{bmatrix}, T(\tau) = \begin{bmatrix} 0 & \frac{\tau}{10} & 0 \\ \frac{7\tau}{10} & 0 & 0 \\ \tau & \frac{\tau}{6} & 0 \end{bmatrix}$$

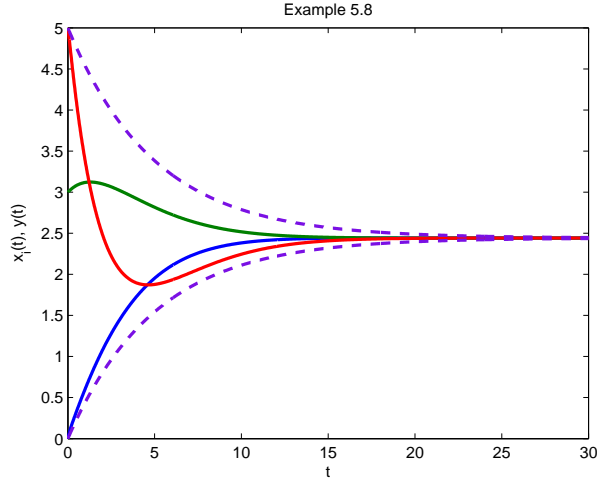


FIGURE 1. Numerical investigations for Example 5.8. The initial conditions are $\phi_1(t) = \sin(8t) + 3t$, $\phi_2(t) = 3\sin(800t) + 3$, $\phi_3(t) = \sin(4t) + 5$. The rate function $y(t) = |y_0|e^{-0.2t}$. The simulations was done with the use of the `ddesd` function in MATLAB.

Then

$$e^{-Lt}L = \begin{bmatrix} 0.2e^{-0.3t} & -0.2e^{-0.3t} & 0 \\ -0.1e^{-0.3t} & 0.1e^{-0.3t} & 0 \\ -1.2e^{-0.4t} + 0.8e^{-0.3t} & 0.8e^{-0.4t} - 0.8e^{-0.3t} & 0.4e^{-0.4t} \end{bmatrix}$$

and

$$G(\gamma, \tau) = \begin{bmatrix} \frac{0.2}{0.3-\gamma} & \frac{0.2}{0.3-\gamma} & 0 \\ \frac{0.1}{0.3-\gamma} & \frac{0.1}{0.3-\gamma} & 0 \\ \frac{0.8}{0.3-\gamma} - \frac{0.4}{0.4-\gamma} & \frac{0.8}{0.3-\gamma} - \frac{0.8}{0.4-\gamma} & \frac{0.4}{0.4-\gamma} \end{bmatrix} e^{\gamma\tau}.$$

For $\tau = 2, \sigma = 0.4$ the matrix $\mathbf{F}(0.2)$ is calculated to be

$$\mathbf{F}(0.2) = \begin{bmatrix} -0.539 & 0.441 & 0 \\ 0.221 & -0.934 & 0 \\ 6.198 & 0.175 & -0.952 \end{bmatrix}$$

and then we can pick $q_1 = 1, q_2 = 1, q_3 \in (0, 0.145)$ to apply Theorem 3.4. The simulation results are depicted in Fig. 1 (see captions for details).

6. Discussion. In this work, we developed a Lyapunov-free argument to the study the stability of a linear distributed consensus system with multiple delays. Our main goal was to establish explicit estimates on the rate of convergence of the solutions, as functions of the system's parameters.

6.1. The FPT Method. This fixed point theory approach consists of three crucial steps. The expression of the solution form in a non-trivial way that will define the solution operator, the choice of a suitable function space and an associated metric and the selection of an appropriate fixed point theorem. Proving the existence of a fixed point of a solution operator in a metric space with prescribed convergence properties is a de-facto solution of the stability problem.

6.1.1. *The solution operator.* We studied the dynamics of system (1) by combining two forms of its solution. The first form, Eq. (5), is a perturbation of the un-delayed system. This way we exploit the valuable kernel e^{-Lt} that describes the global dynamics of the distributed algorithm. The second one, Eq. (7), characterizes the dynamical behavior of the system in a local manner. It expresses the rate at which each agent converges to the weighted sum of the delayed states of its neighboring agents. This representation illustrates the dissipative averaging of the algorithm but sheds no light on the global dynamics. Due to the processing delays, Eq. (7) is in turn a perturbation with respect to the model with only propagation delays. Combining these representations in the same manner as in the Example 1.1, we obtained a new form of the solutions that includes both of these valuable features.

6.1.2. *Metric Space and Distance Function.* Given the initial conditions ϕ and the system parameters, we considered the unique solution of (1) on $[-\max\{\tau, \sigma\}, \tau]$ and defined a space of functions each member of which agrees on $[-\max\{\tau, \sigma\}, \tau]$. This extension of the solution is due to the way we combined the two solution operators (5) and (7) and has no effect in the study of the stability of solutions. Next, each member of this function space converges exponentially fast to a constant value k according to a weighted metric with weights \mathbf{q} and exponent rate γ . We proved that this pair is, for arbitrary but fixed parameters, a complete metric space in Proposition 4.1. Next, based on the defined solution operator (8), we adapted our metric space so that the operator maps this space into itself. In particular, the rate cannot be faster than the local and global exponential rates d_i and $\Re\{\lambda\}$ respectively. Also due to, Eq. (5), the convergence point has a closed form, explicitly defined from the the system parameters and the initial data. Another point of interest is that only if k is finite, the conclusion that our metric space is complete and this is guaranteed by Assumption 3.2. An assumption that in fact puts an upper bound on the difference between the propagation and processing delays.

We turn our attention now to the metric function ρ as defined in Eq. (10). This is a generalized weighted l_∞ -type metric. The supremum over t is necessary for our space to be complete. We chose to work on the maximum over $i \in [N]$ metric because we are dealing with integral equations of a primarily asymmetric system. These types of metrics are typical in distributed systems (see [11] or [13]). We introduced the weights $q_i > 0$ as a design feature in an attempt to capture the geometry of the state space of solutions as it depends on a_{ij} , σ_i^j and τ_j^i . This idea is motivated by Lyapunov's first method in the stability analysis of the general linear system $\dot{x} = Ax$, where one the investigator is asked to derive the classical Lyapunov matrix P in the Lyapunov equation. This way we essentially transformed the contraction problem into a linear inequality problem and applied existence theorems from Convex Analysis in order to ensure that we can find an appropriate metric to make our operator a contraction one, see Appendix C. Example 5.1 clearly illustrates this point.

6.1.3. *Fixed Point Theorem.* The linear nature of the problem we chose the simplest existence and uniqueness theorem, the Contraction Mapping Principle (B.2) that applies function spaces of great flexibility such as the complete metric spaces.

6.2. **Advantages.** In the pros of this method one may account that we need not deal with the difficult task for finding a Lyapunov function which is a generalized metric. As it is already stated the derivation of a Lyapunov function for these systems is limited to the case of no processing delays or with increased connectivity.

The interested reader may consult [24] where an attempt to derive a Lyapunov function similar to Example 5.1 is made. In either case one can hope only for convergence results. This is many real world phenomena is not satisfactory as it is of great importance the question of what is the rate of convergence and how is this affected by the various delays.

The derivation of the solution operator is a metric-free process. Metrics are used in the last but crucial step of proving that the operator is a contraction and they are of course of utmost importance. The whole process allows for the investigator to attain an overall control of the system dynamic behavior and in every step, fruitful information are extracted such as the consensus point and the rate of convergence. The latter one is a function that depends explicitly on the systems parameters. This question cannot be addressed by conventional Lyapunov or frequency methods.

6.3. Disadvantages. FPT methods appear to give a lot information and are able to handle distributed systems in utmost generality; yet at the same time these methods ask for a lot, in terms of computations and analysis. For the sake of justice, we should also mention a number of drawbacks on account of this approach. The points made in this subsection obviously outline the framework of the future work on this problem.

At first one notices that it is a lengthy method. It requires several different steps each of which involves tedious algebraic calculations. The deeper is one willing to dig to sharpen their results, the more calculation they are forced to make. In particular, the process of deriving the contraction property can be very painful exactly because the method will take any supplementary information the investigator is willing to provide. This difficulty lies in the heart of distributed algorithms exactly because a global governing kernel does not exist, other than this of the completely undelayed linear time invariant model which Algebraic Graph Theory provides.

Next, although our result is stated in utmost generality by allowing fully distributed delays and weights under the regime of minimal connectivity, we encounter serious difficulties in proving delay independence results when we neglect processing delays. There is a number of different factors which contribute to this problem. The first factor comes from the solution operator. We effectively considered Eq. (1) as a perturbation to the original un-delayed system. It is only reasonable then to expect delay dependent results (see also [26]). On the other hand, the dimensionality of the problem forced us to make use of the l_∞ norm. Under this norm one encounters the overall dynamics from the point of view of the local dynamical behavior of an arbitrary agent. This is a serious defect which we managed to partially remedy by introducing the weights \mathbf{q} . We tested the results of this approach in the Examples (5.1),(5.5) and (5.6) where we managed to obtain semi-delay independent stability. Unfortunately, \mathbf{q} imposes additional computational complexity to the problem. It is intuitively clear, however, that the more asymmetrical a system is the more difficult it's analysis becomes. Example, 5.7 clearly illustrates that any such asymmetry must be compensated with smaller and smaller maximum allowed delay.

6.4. Variations and Extensions. We conclude this section by discussing a number of important extensions that can be similarly handled by the theory developed in this paper.

6.4.1. Time-Varying Dynamics. We chose a simple linear time invariant communication topology, already known in literature because it is our purpose to focus on

the effect of delays. Time varying systems can just as easily be handled with FPT methods. While a Lyapunov function would require for the connectivity weights to satisfy $a_{ij}(t) \geq 0$ for every t and $\int_{-\infty}^{\infty} a_{ij}(s)ds = \infty$, FPT arguments require only the latter condition. It is a standard advantage of FPT methods over Lyapunov, where only averaging conditions are needed and pointwise [3]. In time varying communication schemes one should, similarly, assume that the undelayed governing kernel, which is a principal transition matrix, converges to a consensus state. For a preliminary introduction to FPT methods and time-varying consensus systems the interested reader is referred to [23]. That model can also include the interesting case of time-switching network topologies such as these discussed in [7] in a unified framework with the case of static time varying topologies. This is because FPT methods are primarily dealing with integral equations which allow these discontinuities in the signals. The only extra consideration is some mild technical consideration on the switching signal and the integration by parts formulae.

6.4.2. *Noise.* State-independent disturbances are a classic tool to model uncertainties in the structure of our system or to model exogenous noise effects. In particular one may choose to study the model

$$\dot{x}_i(t) = \sum_j a_{ij}(x_j(t - \sigma) - x_i(t - \tau)) + f_{ij}(t), t \geq t_0$$

In the study of consensus solutions one can simply ask $f_{ij} \in L^1$. The analysis is then straightforward as the noisy term f_{ij} contributes nothing in the derivation of the contraction condition. In particular for $f_{ij} \in L^1$ the consensus point is defined as

$$k = \frac{\sum_i c_i [\phi_i(0) + \sum_{j \in N_i} a_{ij} (\int_{-\tau_j^i}^0 \phi_j(s)ds - \int_{-\sigma_j^i}^0 \phi_j(s)ds) + \sum_{j \in N_i} \int_{t_0}^{\infty} f_{ij}(s)ds]}{\sum_i c_i [1 + \sum_{j \in N_i} a_{ij} (\tau_j^i - \sigma_j^i)]}$$

Furthermore, f_{ij} can also be T -periodic functions of time and then one can prove consensus to a periodic solution (synchronization) by carefully adapting the metric space. We also comment that f_{ij} can be stochastic perturbations and such modeling approach is considered for a non-linear model in [22].

6.4.3. *Distributed Delays.* Uncertainties in the delays are typically modeled by considering distributions instead of constant values τ, σ . In [26] we considered uncertainty in the delays of a simplified version of (1)

$$\dot{x}_i(t) = \sum_j a_{ij} \int_{t-\tau}^t p_{ij}(s-t)x_j(s)ds - \sum_j a_{ij}x_i(t) \quad (16)$$

with $a_{ij} = a_{ji}$ and for some distribution functions $\int_{-\tau}^0 p_{ij}(s)ds = 1$ for all $j \in N_i$. The result we derived is:

Theorem 6.1. [26] Consider the sum $\tilde{B} = \sum_{i=1}^N \sum_{j \in N_i} \tilde{b}_{ij}$ where \tilde{b}_{ij} are defined by

$$\tilde{b}_{ij} = \begin{cases} 0 & \text{if } j \notin N_i \\ a_{ij} \int_{-\tau}^0 |p_{ij}(s)| (e^{-ds} - 1) ds & \text{if } j \in N_i \end{cases}$$

Under Assumption 3.1 and $a_{ij} = a_{ji}$, if

$$\tilde{B} \frac{e^{d\tau} - 1}{d} \left(1 + \frac{\sqrt{N} \|L\|}{\lambda_2 - \gamma} \right) < 1$$

then the solutions of (16) converge to

$$k_{(16)} := \frac{\sum_i \phi_i(0) + \sum_{i,j} a_{ij} \int_{-\tau}^0 p_{ij}(s) \int_s^0 \phi_j(w) dw ds}{N + \sum_{i,j} a_{ij} \int_{-\tau}^0 p_{ij}(s)(-s) ds}$$

exponentially fast with rate γ . Note the similarities of the contraction condition of the above result with Eq. (14).

6.4.4. Non-Linearities. Non-Linearities are as inevitable as are delays and within the context of the model in question they can appear either in the connectivity weights, the delayed arguments or even the noise components. Such mathematical models are beyond the scopes of this work and thus we will only make a few comments. Despite the lengthy process of deriving the solution operator, the linearity of the problem allowed us to work on complete metric spaces. Non-linear systems would require strong Lipschitz conditions, otherwise one would seek to prove existence of fixed points in compact linear spaces using more involved fixed point theorems such as these of Schauder or Krasnoselskii [21, 3]; in which case, however, crucial features such as uniqueness are to be sacrificed. For a treatment of non-linear distributed consensus systems with delays and FPT methods by means of the Contraction Mapping Principle the interested reader is referred to [22] and [25].

7. Conclusions and Future Work. In this work, we examined a delayed and perturbed version of distributed continuous time consensus algorithms using classical Fixed Point Theory arguments. We explained why Lyapunov-based methods can easily become very complex or of limited applicability in such types of systems. Using the Contraction Mapping Principle, we established sufficient delay-dependent conditions for asymptotic stability to a consensus state and at the same time we approximated the rate of convergence. Although in the existing literature, delay-independent asymptotic results exist, these results treat the case of negligible processing delays. In addition, none of these results calculated the rate of convergence, which is very important in real-life applications. We showed that Fixed Point Theory can handle such cases with remarkable flexibility. A number of important steps are to be made among which the most imminent one is to improve the results to become definitively delay-independent in the case of no processing delays with the accompanying estimate on the rate of convergence.

Appendix A. Algebraic Graph Theory and Agreement Dynamics [5, 2, 10].

The mathematical object that is most widely used to model the communication structure among N agents is the weighted directed graph. This is defined as the triple $G = ([N], E, W)$ where $[N]$ is the set of nodes, E is a subset of $[N] \times [N]$ which defines the established communication connections among nodes and W is a set of non-negative numbers so that $a_{ij} \in W$ models the strength of the connection from node j to node i , with the convention that $a_{ij} = 0$ if and only if $(j, i) \notin E$. This is a directed graph and in this work we will be interested in sufficiently connected graphs, i.e. graphs that are routed-out branching. A routed out branching graph is a graph that contains a spanning tree (i.e. there exists a node in $[N]$ from which one can reach any other node). Given E , each agent i has a neighbourhood of nodes, to which it is adjacent. We denote by N_i the subset of $[N]$ such that $(j, i) \in E$. In the routed-out branching graph there can be at most one node, named i , with $d_i = 0$. This is called the *leader* of G as it is not affected by anyone while it affects eventually the rest of the nodes. If all nodes are affected by

others the communication topology is called as *leaderless*. The notation $\sum_{i,j}$ stands for $\sum_{i=1}^N \sum_{j \in N_i}$. A matrix representation of G is through the adjacency matrix $A = [a_{ij}]$, the degree matrix $D = \text{Diag}[d_i]$, $d_i := \sum_{j \in N_i} a_{ij}$ and the Laplacian $L := D - A$. We will also use the symbol $d_i^+ = \min_{i \in [N]} \{d_i : d_i > 0\}$. The spectrum of L is denoted by $\{\lambda_1, \lambda_2, \dots, \lambda_N\}$ and we set $\Re(\lambda) := \min_{i \geq 2} \{\Re(\lambda_i)\}$. Let also \mathbf{c} be the left eigenvector of L normalized so that $\mathbf{c}^T \mathbb{1} = 1$.

A.1. Agreement Dynamics. We consider the initial value problem

$$\dot{\mathbf{x}} = -L\mathbf{x}, \quad \mathbf{x}(0) = \mathbf{x}_0 \quad (17)$$

It's dynamics can be analysed by standard linear algebra tools (see for example [8]), much of which is summarized in the following result.

Proposition A.1. *Let G be a routed-out branching graph and denote by L it's Laplacian matrix. The following properties hold:*

1. $L\mathbf{p} = 0$ if and only if $\mathbf{p} \in \Delta$.
2. The spectrum of L , $\{\lambda_i\}_{i=1}^N$, can be enumerated so that $\lambda_1 = 0$ and $\Re(\lambda_i) > 0$, $i \geq 2$.
3. The left eigenvector of L , \mathbf{c} is unique, up to normalization and non-negative elementwise.
4. There exists $J > 0 : \|e^{-Lt} - \mathbb{1}\mathbf{c}^T\| \leq J e^{-\Re(\lambda)t}$.
5. $e^{-Lt}L$ is a Laplacian matrix which asymptotically vanishes exponentially fast with rate no worse than $\Re(\lambda)$.

Proof. For (1), (2) see Propositions 3.8 and 3.11 of [10] respectively. For (3) we work as follows: from (2) we have that $\text{rank}(L) = N - 1$. Then if $\mathbf{c}_1, \mathbf{c}_2$ are two left eigenvectors of L associated with the zero eigenvalue then $\mathbf{c}_1 = \epsilon \mathbf{c}_2$ for some $\epsilon > 0$. From the normalization condition $\mathbf{c}_1^T \mathbb{1} = \epsilon \mathbf{c}_2^T \mathbb{1} = 1$ so that in fact $\mathbf{c}_1 = \mathbf{c}_2$ and uniqueness follows. To conclude it suffices to show that the elements of \mathbf{c} are of the same sign. Assume, for the sake of contradiction that the result does not hold and without loss of generality let c_1, c_2, \dots, c_r be the negative components of \mathbf{c} . Then $\mathbf{c}^T \mathbb{1} = 1$ implies $r \leq N - 1$. Next, $\mathbf{c}^T L = 0$ gives

$$\sum_{j=1}^N a_{ij} c_i = \sum_{j=1}^r a_{ji} c_j + \sum_{j=r}^N a_{ji} c_j, \quad i = 1, \dots, N$$

Take the sum of the first r equations and after cancelling the common terms observe that the resulting equation has negative left hand-side and positive right hand side, a contradiction. The zero terms of \mathbf{c} were neglected as they play no role in the proof. For (4), consider the Jordan canonical form of $L := PJ(L)P^{-1}$ so that $e^{-Lt} = P e^{-J(L)t} P^{-1}$. For the Jordan blocks, we know that $J(\lambda_1) = J(0) = 0$ so that both $LP = PJ$ and $P^{-1}L = JP^{-1}$ have a zero first row. So the first row of P^{-1} is the left eigenvector \mathbf{c} , canonicalized so that $\mathbf{c}^T \mathbb{1} = 1$ and the first column of L is the right eigenvector of L chosen to be $\mathbb{1}$. The projection of any vector ζ onto Δ is $\mathbb{1}\mathbf{c}^T \zeta$. Let L have $q \leq N$ distinct eigenvalues. Again $\lambda_1 = 0$ and $\Re\{\lambda_i\} > 0$ for $i \geq 2$. Since any $\zeta \in \mathbb{C}^N$ can be written as $\zeta = \mathbf{w}_1 + \dots + \mathbf{w}_q$ where $\mathbf{w}_j \in M(\lambda_j)$, the generalized eigenspace for λ_j and in particular, $\mathbf{w}_1 = \mathbb{1}\mathbf{c}^T \zeta$. Then by standard calculations

$$e^{-Lt} \zeta = \mathbb{1}\mathbf{c}^T \zeta + \sum_{j=2}^q e^{-\lambda_j t} \sum_{k=0}^{r(-\lambda_j)-1} (-L + \lambda_j I)^k \frac{t^k}{k!} \mathbf{w}_j \quad (18)$$

and the rest of the proof follows exactly as in Theorem 5.8 of [8].

For the last one, observe that $e^{-Lt}L$ is a Laplacian by the definition of the exponential of a matrix and the fact that any power of a Laplacian matrix is Laplacian and whereas the sum of two Laplacian matrices is a Laplacian matrix. Finally since $e^{-Lt}L = (e^{-Lt} - \mathbb{1}\mathbf{c}^T)L$ we get that that $\|Le^{-Lt}\| \leq \|L\|Je^{-\Re(\lambda)t}$. \square

Remark A.2. In case of an undirected network the Laplacian is a symmetric positive semi-definite matrix with real spectrum: $\{0 < \lambda_2 \leq \dots \leq \lambda_N\}$ making the analysis significantly simpler (see [24, 26]). Then, $\mathbf{c} = \mathbb{1}\frac{1}{N}$ and the convergence to $\mathbf{c}^T \mathbf{x}_0$ is exponential with rate λ_2 .

Remark A.3. From Proposition A.1 we can conclude on the form $e^{-Lt}L = [\kappa_{ij}(t)]_{ij}$ where $\sum_{j \neq i} \kappa_{ij} = -\kappa_{ii}$ and $\kappa_{ij}(t) \rightarrow 0$ exponentially fast. These functions are calculated from the connectivity weights and are assumed known.

Appendix B. Metric Spaces and Fixed Point Theory. An arbitrary metric space is defined axiomatically.

Definition B.1. A pair (\mathbb{M}, ρ) is a metric space if \mathbb{M} is a set and $\rho : \mathbb{M} \times \mathbb{M} \rightarrow [0, \infty)$ such that when x, y, z are in \mathbb{M} then

- $\rho(x, z) \geq 0$, $\rho(y, y) = 0$ and $\rho(y, z) = 0$ implies $y = z$.
- $\rho(y, z) = \rho(z, y)$ and
- $\rho(y, z) \leq \rho(y, x) + \rho(x, z)$

Proof of Proposition 4.1. It is trivial to show that ρ as defined in Eq. (10) is a metric function according to Definition (B.1). It then suffices to show that a Cauchy sequence in \mathbb{M} has a limit in \mathbb{M} . Let $\{\phi_i\}$ be such a sequence. Then

$$\|\phi_j(t) - \phi_i(t)\| \leq e^{\gamma t} \|\phi_j(t) - \phi_i(t)\| \leq \rho(\phi_i, \phi_j)$$

implies that $\phi_j(t)$ is a Cauchy sequence in $(\mathbb{R}^N, \|\cdot\|)$ for all t , so $\phi_j(t) \rightarrow \phi(t)$. We will show that $\phi \in \mathbb{M}$ and this is the result of the following claims. The first claim is to prove that ϕ is continuous: Given $\varepsilon > 0$ there exists Q such that $\sup_{t \geq t_0} \|\phi_i(t) - \phi_j(t)\| < \varepsilon$ for $j, i > Q$. Fix such $j > Q$ and let $i \rightarrow \infty$ from above we get $\|\phi(t) - \phi_j(t)\| < \varepsilon$ for all t . So $\phi_k \rightrightarrows \phi$ and thus ϕ is continuous.⁴ The second claim is that ϕ is bounded. Indeed,

$$\|\phi(t)\| \leq \sup_t \|\phi(t) - \phi_j(t)\| + \sup_t \|\phi_j(t)\| < \infty$$

The third claim is that $\phi \rightarrow \Delta$: For any $\varepsilon > 0$ take $Q > 0$ such that $j > Q$ implies $\|\phi(t) - \phi_j(t)\| < \frac{\varepsilon}{2}$. Fix $j > Q$ and $t > T$ such that $\sup_{t \geq T} \|\phi(t) - \phi_j(t)\| < \frac{\varepsilon}{2}$ for $j > Q$ and $\|\phi_j(t) - \lim_t \phi_j(t)\| < \frac{\varepsilon}{2}$ for $t > T$. Then

$$\|\phi(t) - \lim_t \phi_j(t)\| \leq \|\phi(t) - \phi_j(t)\| + \|\phi_j(t) - \lim_t \phi_j(t)\| < \varepsilon$$

Finally for any $R > 0$ pick i, j large enough so that $\rho(\phi_i, \phi_j) < R$ which implies that $\|\phi_i(t) - \phi_j(t)\| \leq Re^{-\gamma t}$ and taking the limit for i , $\|\phi(t) - \phi_j(t)\| \leq Re^{-\gamma t}$ for all t . Then

$$\|\phi(t) - \mathbb{1}k\| \leq \|\phi_j(t) - \mathbb{1}k\| + \|\phi_j(t) - \phi(t)\| \leq (R_j + R)e^{-\gamma t}$$

so that $\sup_t e^{\gamma t} \|\phi(t) - \mathbb{1}k\| < \infty$. \square

⁴the symbol \rightrightarrows stands for uniform convergence.

B.1. The Contraction Mapping Principle. Given two metric spaces (\mathbb{M}_i, ρ_i) for $i = 1, 2$ an operator $\mathcal{P} : \mathbb{M}_1 \rightarrow \mathbb{M}_2$ is a *contraction* if there exists a constant $\alpha \in [0, 1)$ such that $\forall x_1, x_2 \in \mathbb{M}_1$

$$\rho_2(\mathcal{P}x_1, \mathcal{P}x_2) \leq \alpha \rho_1(x_1, x_2)$$

The next celebrated theorem, is the Contraction Mapping Principle and its proof can be found in any advanced analysis or ordinary differential equations textbook [20, 3, 8].

Theorem B.2. *Let (\mathbb{M}, ρ) be a complete metric space and $\mathcal{P} : \mathbb{M} \rightarrow \mathbb{M}$ a contraction operator. Then there is a unique $x \in \mathbb{M}$ such that $\mathcal{P}x = x$.*

Appendix C. Theory of Linear Inequalities. We will need the following result from [19], Section 22.

Theorem C.1. *Let $\mathbf{a}_i \in \mathbb{R}^m$ and $\alpha_i \in \mathbb{R}$ for $i = 1, \dots, m$ and let k be an integer, $1 \leq k \leq m$. Assume that the system*

$$\mathbf{a}_i^T \boldsymbol{\xi} \leq \alpha_i$$

for $i = k + 1, \dots, m$ is consistent. Then one and only one of the following alternatives hold:

1. *There exists a vector $\boldsymbol{\xi} \in \mathbb{R}^n$ such that*

$$\begin{aligned} \mathbf{a}_i^T \boldsymbol{\xi} &< \alpha_i, & i = 1, \dots, k \\ \mathbf{a}_i^T \boldsymbol{\xi} &\leq \alpha_i, & i = k + 1, \dots, m \end{aligned}$$

2. *There exist non-negative real numbers $\zeta_i |_{i=1}^m$ such that at least one of $\zeta_i |_{i=1}^k$ is not zero and*

$$\sum_{i=1}^m \zeta_i \mathbf{a}_i = 0 \quad \text{and} \quad \sum_{i=1}^m \zeta_i \alpha_i \leq 0.$$

Appendix D. Proofs. From the definition of \mathcal{P} in Eq. (11) we observe that for $t \geq \tau$ it can be written as the sum

$$\mathcal{P} = \mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4 \tag{19}$$

where

$$\mathcal{P}_1(t) := e^{-D\tau} e^{-L(t-\tau)} \mathbf{r}_0$$

$$\mathcal{P}_2(t) := \sum_{i,j} \int_{t-\tau-\tau_j^i}^{t-\tau} \left(e^{-D(t-\tau_j^i-s)} - e^{-D\tau} \right) A_{ij} \mathbf{x}(s) ds + \sum_{i,j} \int_{t-\tau}^{t-\tau_j^i} e^{-D(t-\tau_j^i-s)} A_{ij} \mathbf{x}(s) ds$$

$$\mathcal{P}_3(t) := \sum_{i,j} B_{ij} \int_{t-\sigma_i^j}^t \mathbf{x}(s) ds - \int_{t-\tau}^t e^{-D(t-s)} D B_{ij} \int_{s-\sigma_i^j}^s \mathbf{x}(w) dw ds$$

$$\mathcal{P}_4(t) := e^{-D\tau} \int_0^{t-\tau} e^{-L(t-\tau-s)} L \sum_{i,j} \left[A_{ij} \int_{s-\tau_j^i}^s \mathbf{x}(w) dw - B_{ij} \int_{s-\sigma_i^j}^s \mathbf{x}(w) dw \right] ds$$

We proceed with the proofs of Proposition 4.2 and Lemma 4.3.

Proof of Proposition 4.2. The proof consists of two parts. The first one the calculation k and the second is the estimate of γ . For the first part we will need the following result

Lemma D.1. *Let L be the weighted Laplacian matrix of a routed-out branching graph G with \mathbf{c} it's normalized left eigenvector as defined in Proposition A.1. Let $\mathbf{z} \in C^1([0, \infty), \mathbb{R}^N)$ such that $\lim_{t \rightarrow \infty} \mathbf{z}(t) \in \mathbb{R}^N$. Then*

$$\lim_{t \rightarrow \infty} \int_0^t e^{-L(t-s)} L \mathbf{z}(s) ds = (I - \mathbb{1} \mathbf{c}^T) \mathbf{z}(\infty)$$

We will calculate the t limit of $(\mathcal{P}\mathbf{x})(t)$ as the sum of the four quantities defined in (19)

For $x \in \mathbb{M}$, $\lim_{t \rightarrow \infty} \mathcal{P}_i(t)$ yields

$$\begin{aligned} \lim_t \mathcal{P}_1(t) &= \lim_t e^{-D\tau} (e^{-L(t-\tau)} - \mathbb{1} \mathbf{c}^T) \mathbf{r}_0 + e^{-D\tau} \mathbb{1} \mathbf{c}^T \mathbf{r}_0 = e^{-D\tau} \mathbb{1} \mathbf{c}^T \mathbf{r}_0 \\ \lim_t \mathcal{P}_2(t) &= e^{-D\tau} \left(\sum_{j \in N_1} \frac{a_{1j}}{d_1} (e^{d_1 \tau_j^1} - 1), \dots, \sum_{j \in N_N} \frac{a_{Nj}}{d_N} (e^{d_N \tau_j^N} - 1) \right)^T k \\ &\quad - e^{-D\tau} \left(\sum_{j \in N_1} a_{1j} \tau_j^1, \dots, \sum_{j \in N_N} a_{Nj} \tau_j^N \right)^T k \\ &\quad + \left(\mathbb{1} - e^{-D\tau} \left(\sum_{j \in N_1} \frac{a_{1j}}{d_1} e^{d_1 \tau_j^1}, \dots, \sum_{j \in N_N} \frac{a_{Nj}}{d_N} e^{d_N \tau_j^N} \right)^T \right) k \\ \lim_t \mathcal{P}_3(t) &= e^{-D\tau} \left(\sum_{j \in N_1} a_{1j} \sigma_1^j, \dots, \sum_{j \in N_N} a_{Nj} \sigma_N^j \right)^T k \end{aligned}$$

Finally, from Lemma D.1 with $\mathbf{z}(t) = \sum_{i,j} \left[A_{ij} \int_{t-\tau_j^i}^t \mathbf{x}(s) ds - B_{ij} \int_{t-\sigma_i^j}^t \mathbf{x}(s) ds \right]$ we obtain

$$\lim_t \mathcal{P}_4(t) = e^{-D\tau} (I - \mathbb{1} \mathbf{c}^T) \left(\sum_{j \in N_1} a_{ij} (\tau_j^1 - \sigma_1^j), \dots, \sum_{j \in N_N} a_{Nj} (\tau_j^N - \sigma_N^j) \right)^T k$$

Take $\mathbf{w} = \left(\sum_{j \in N_1} a_{ij} (\tau_j^1 - \sigma_1^j), \dots, \sum_{j \in N_N} a_{Nj} (\tau_j^N - \sigma_N^j) \right)^T$ and cancel the common terms to obtain

$$\lim_t (\mathcal{P}\mathbf{x})(t) = \mathbb{1} k + e^{-D\tau} \mathbb{1} \left(\mathbf{c}^T \mathbf{r}_0 - k - \mathbf{c}^T \mathbf{w} k \right) = \mathbb{1} k$$

if k is defined as in (4). For the second part, Proposition A.1 implies $\| (e^{-Lt} - \mathbb{1} \mathbf{c}^T) \mathbf{r}_0 \| = \mathcal{O}(t^{r(\lambda)-1} e^{-\Re(\lambda)t})$ whereas for $\mathbf{x} \in \mathbb{M}_{\gamma,k}$ the rest of the terms of $(\mathcal{P}\mathbf{x})(t)$ are of order $\mathcal{O}(e^{\gamma t})$. Then $\sup_{t \geq p} e^{\gamma t} \| (\mathcal{P}\mathbf{x})(t) - \mathbb{1} k \|$ is finite for any $\gamma < \Re(\lambda)$. \square

Proof of Lemma D.1. Set $\mathbf{z}^c(t) := \mathbf{z}(t) - \mathbb{1} \mathbf{c}^T \mathbf{z}(t)$. Then

$$\begin{aligned} Q(t) &:= \int_0^t e^{-L(t-s)} L \mathbf{z}(s) ds = \int_0^t (e^{-L(t-s)} - \mathbb{1} \mathbf{c}^T) L \mathbf{z}^c(s) ds \\ &= \int_0^t \frac{d}{ds} (e^{-L(t-s)} - \mathbb{1} \mathbf{c}^T) \mathbf{z}^c(s) ds = \int_0^t d(e^{-L(t-s)}) \mathbf{z}^c(s) ds \\ &= \mathbf{z}^c(t) - (e^{-Lt} - \mathbb{1} \mathbf{c}^T) \mathbf{z}^c(0) - \int_0^t (e^{-L(t-s)} - \mathbb{1} \mathbf{c}^T) \dot{\mathbf{z}}^c(s) ds \end{aligned}$$

The result follows from Assumption 3.1 and hence the observation that the integral vanishes it is a convolution of an L^1 function with a function that tends to zero. \square

Proof of Lemma 4.3. Take $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{M}$. Then

$$\begin{aligned} \rho((\mathcal{P}\mathbf{x}_1), (\mathcal{P}\mathbf{x}_2)) &= \sup_{t \geq \tau} e^{\gamma t} \|(\mathcal{P}\mathbf{x}_1)(t) - (\mathcal{P}\mathbf{x}_2)(t)\|_{\mathbf{q}} \\ &\leq \sum_{l=1}^4 \sup_{t \geq \tau} e^{\gamma t} \|(\mathcal{P}_l \mathbf{x}_1)(t) - (\mathcal{P}_l \mathbf{x}_2)(t)\|_{\mathbf{q}} \end{aligned}$$

The contribution of each $\mathcal{P}_l \mathbf{x}_1 - \mathcal{P}_l \mathbf{x}_2$ is studied separately. At first, $\mathcal{P}_1 \mathbf{x}_1 - \mathcal{P}_1 \mathbf{x}_2$ contributes nothing. For the rest we work as follows.

\mathcal{P}_2 . We estimate the upper bound of $e^{\gamma t} q_i |(\mathcal{P}_2 \mathbf{x}_1)^{(i)}(t) - (\mathcal{P}_2 \mathbf{x}_2)^{(i)}(t)|$. Observe that $e^{-d_i(t-s-\tau_j^i)} - e^{-d_i \tau}$ is non-negative for $s \in [t - \tau - \tau_j^i, t - \tau]$ and for convenience set $\mathbf{x}_{12}(s) := \mathbf{x}_1(s) - \mathbf{x}_2(s)$ and $\rho := \sup_{t \geq \tau} e^{\gamma t} \|\mathbf{x}_{12}(t)\|_{\mathbf{q}}$.

$$\begin{aligned} &\left| \int_{t-\tau-\tau_j^i}^{t-\tau} (e^{-d_i(t-s-\tau_j^i)} - e^{-d_i \tau}) a_{ij} \mathbf{x}_{12}^{(j)}(s) ds + \int_{t-\tau}^{t-\tau_j^i} e^{-d_i(t-s-\tau_j^i)} a_{ij} \mathbf{x}_{12}^{(j)}(s) ds \right| \leq \\ &\leq \left[\int_{t-\tau-\tau_j^i}^{t-\tau} (e^{-d_i(t-s-\tau_j^i)} - e^{-d_i \tau}) a_{ij} e^{-\gamma s} ds + \int_{t-\tau}^{t-\tau_j^i} e^{-d_i(t-s-\tau_j^i)} a_{ij} e^{-\gamma s} ds \right] \frac{1}{q_j} \rho \end{aligned}$$

\Rightarrow where the two integrals give

$$\begin{aligned} e^{\gamma t} \int_{t-\tau-\tau_j^i}^{t-\tau} (e^{-d_i(t-s-\tau_j^i)} - e^{-d_i \tau}) a_{ij} e^{-\gamma s} ds &= \frac{a_{ij}}{d_i - \gamma} \left(e^{d_i(\tau_j^i - \tau) + \gamma \tau} - e^{-(d_i - \gamma)\tau + \gamma \tau_j^i} \right) - \\ &- \frac{a_{ij}}{\gamma} e^{-(d_i - \gamma)\tau} (e^{\gamma \tau_j^i} - 1) \\ e^{\gamma t} \int_{t-\tau}^{t-\tau_j^i} e^{-d_i(t-s-\tau_j^i)} a_{ij} e^{-\gamma s} ds &= \frac{a_{ij}}{d_i - \gamma} \left(e^{\gamma \tau_j^i} - e^{d_i(\tau_j^i - \tau) + \gamma \tau} \right) \end{aligned}$$

Check that the first and last term cancel and that summing over $j \in N_i$ we obtain the following estimate

$$q_i \sum_{j \in N_i} \left[a_{ij} e^{\gamma \tau_j^i} \frac{1 - e^{-(d_i - \gamma)\tau}}{d_i - \gamma} - a_{ij} e^{-(d_i - \gamma)\tau} \frac{e^{\gamma \tau_j^i} - 1}{\gamma} \right] \frac{1}{q_j} \rho \quad (20)$$

Remark D.2. As $\gamma \downarrow 0$, the last expression becomes

$$q_i \sum_{j \in N_i} a_{ij} \left[\frac{1 - e^{-d_i \tau}}{d_i} - \tau_j^i e^{-d_i \tau} \right] \frac{1}{q_j} \rho$$

\mathcal{P}_3 . Similar manipulation as above yields

$$\begin{aligned} &e^{\gamma t} \left| \sum_{j \in N_i} a_{ij} \int_{t-\sigma_i^j}^t \mathbf{x}_{12}^{(i)}(s) ds \right| + e^{\gamma t} \left| \sum_{j \in N_i} a_{ij} \int_{t-\tau}^t d_i e^{-d_i(t-s)} \int_{s-\sigma_i^j}^s \mathbf{x}_{12}^{(i)}(s) ds \right| \\ &\leq \sum_{j \in N_i} a_{ij} \frac{e^{\gamma \sigma_i^j} - 1}{\gamma} \rho + \sum_{j \in N_i} a_{ij} \frac{e^{\gamma \sigma_i^j} - 1}{\gamma} d_i \frac{1 - e^{-(d_i - \gamma)\tau}}{d_i - \gamma} \rho \end{aligned}$$

Note here that in this case the weights q_i 's are canceled and so we get the estimate

$$\sum_{j \in N_i} a_{ij} \frac{e^{\gamma \sigma_i^j} - 1}{\gamma} \left(1 + d_i \frac{1 - e^{-(d_i - \gamma)\tau}}{d_i - \gamma} \right) \rho \quad (21)$$

Remark D.3. As $\gamma \downarrow 0$, the last expression becomes

$$\sum_{j \in N_i} a_{ij} \sigma_i^j (2 - e^{-d_i \tau}) \rho$$

\mathcal{P}_4 . Finally,

$$\mathcal{P}_4(t) := e^{-D\tau} \int_0^{t-\tau} e^{-L(t-\tau-s)} L \sum_{l,m} \left[A_{lm} \int_{s-\tau_m^l}^s \mathbf{x}_{12}(w) dw - B_{lm} \int_{s-\sigma_l^m}^s \mathbf{x}_{12}(w) dw \right] ds$$

Let κ_{ij} be the $(i,j)^{th}$ element of $e^{-Lt}L$ so that $\kappa_{ii} = -\sum_{j \neq i} \kappa_{ij}$. A careful calculation for on the i^{th} row of $e^{-Lt}L \sum_{l,m} A_{lm} \int_{s-\tau_m^l}^s \mathbf{x}_{12}(w) dw$ yields

$$\begin{aligned} & \sum_{l=1}^N \kappa_{il} \sum_{j=1}^N a_{lj} \int_{s-\tau_j^l}^s \mathbf{x}_{12}^{(j)}(w) dw = \\ & \kappa_{ii} \sum_{j=1}^N a_{ij} \int_{s-\tau_j^i}^s \mathbf{x}_{12}^{(j)}(w) dw + \sum_{l \neq i} \kappa_{il} \sum_{j=1}^N a_{lj} \int_{s-\tau_j^l}^s \mathbf{x}_{12}^{(j)}(w) dw = \\ & \sum_{l \neq i} \kappa_{il} \left[\sum_{j=1}^N \left(a_{lj} \int_{s-\tau_j^l}^s \mathbf{x}_{12}^{(j)}(w) dw - a_{ij} \int_{s-\tau_j^i}^s \mathbf{x}_{12}^{(j)}(w) dw \right) \right] = \\ & \sum_{l \neq i} \kappa_{il} \left[\sum_{j=1}^N \left((a_{lj} - a_{ij}) \int_{s-\tau_j^l}^s \mathbf{x}_{12}^{(j)}(w) dw + a_{ij} \int_{s-\max\{\tau_j^l, \tau_j^i\}}^{s-\min\{\tau_j^l, \tau_j^i\}} \mathbf{x}_{12}^{(j)}(w) dw \right) \right] \end{aligned}$$

Recall the notations g_{il} and $h_{l,j,i}$ as in Eq. (2) and (3) respectively. Then the first bound is

$$q_i e^{-d_i \tau} \sum_{l \neq i} \sum_{j=1}^N h_{l,j,i}(\gamma) g_{il}(\gamma) \frac{1}{q_j} \rho$$

The second bound is

$$q_i e^{-d_i \tau} \sum_{l=1}^N \sum_{j=1}^N g_{il}(\gamma) a_{lj} \frac{e^{\gamma \sigma_l^j} - 1}{\gamma} \frac{1}{q_l} \rho$$

We add them both to obtain

$$q_i e^{-d_i \tau} \left[\sum_{l \neq i} \sum_{j=1}^N h_{l,j,i}(\gamma) g_{il}(\gamma) \frac{1}{q_j} + \sum_{l=1}^N \sum_{j=1}^N g_{il}(\gamma) a_{lj} \frac{e^{\gamma \sigma_l^j} - 1}{\gamma} \frac{1}{q_l} \right] \rho \quad (22)$$

Remark D.4. As $\gamma \downarrow 0$, the last expression becomes

$$q_i e^{-d_i \tau} \left[\sum_{l \neq i} \sum_{j=1}^N h_{l,j,i}(0) g_{il}(0) \frac{1}{q_j} + \sum_{l=1}^N \sum_{j=1}^N g_{il}(0) a_{lj} \sigma_l^j \frac{1}{q_l} \right] \rho$$

Combine Remarks D.2, D.3, D.4 and reorder the weights q_i to obtain the condition of Assumption 3.3. □

REFERENCES

- [1] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, pages 2996 – 3000, Dec. 2005.
- [2] B Bollobas. *Modern Graph Theory*, volume 184. Springer, 1998.
- [3] T.A. Burton. *Stability by Fixed Point Theory for Functional Differential Equations*. Dover Books on Mathematics Series. Dover Publications, 2006.
- [4] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Transactions on Automatic Control*, 52(5):852–862, 2007.
- [5] C. Godsil and G. Royle. *Algebraic Graph Theory*, volume 207. Springer, 2001.

- [6] J. K. Hale and S. M. Verduyn Lunel. *Introduction to functional-differential equations*, volume 99 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1993.
- [7] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003.
- [8] N. G. Markley. *Principles of Differential Equations*. Pure and Applied Mathematics. Wiley-Interscience Series of Texts, Monographs, and Tracts, New Jersey, 2004.
- [9] I. Matei, N. Martins, and J. S. Baras. Almost sure convergence to consensus in markovian random graphs. *2008 47th IEEE Conference on Decision and Control*, pages 3535–3540, 2008.
- [10] M. Mesbahi and M. Egerstedt. *Graph Theoretic Methods in Multiagent Networks*. Princeton University Press, 2010.
- [11] L. Moreau. Stability of continuous-time distributed consensus algorithms. *43rd IEEE Conference on Decision and Control*, page 21, 2004.
- [12] S. Motsch and E. Tadmor. A new model for self-organized dynamics and its flocking behavior. *Journal of Statistical Physics*, 144(5):923–947, 2011.
- [13] S. Motsch and E. Tadmor. Heterophilous dynamics enhances consensus. *preprint*, Jan. 2013.
- [14] U. Munz, A. Papachristodoulou, and F. Allgower. Delay-dependent rendezvous and flocking of large scale multi-agent systems with communication delays. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 2038 –2043, dec. 2008.
- [15] R. Olfati-Saber and R.M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *Automatic Control, IEEE Transactions on*, 49(9):1520 – 1533, sept. 2004.
- [16] A. Papachristodoulou, A. Jadbabaie, and U. Munz. Effects of delay in multi-agent consensus and oscillator synchronization. *Automatic Control, IEEE Transactions on*, 55(6):1471 –1477, june 2010.
- [17] M. Porfiri and D. L. Stilwell. Consensus seeking over random weighted directed graphs. *Automatic Control IEEE Transactions on*, 52(9):1767–1773, 2007.
- [18] E. J. Putzer. Avoiding the jordan canonical form in the discussion of linear systems with constant coefficients. *The American Mathematical Monthly*, 73(1):2–7, 1966.
- [19] R.T. Rockafellar. *Convex Analysis*. Convex Analysis. Princeton University Press, 1997.
- [20] H. L. Royden. *Real Analysis*. McMillan Publishing Company, 3 edition, 1989.
- [21] D.R. Smart. *Fixed Point Theorems*. Cambridge University Press, London, 1980.
- [22] C. Somarakis and J. S. Baras. Asymptotic consensus solutions in non-linear delayed distributed algorithms under deterministic & stochastic perturbations. In *52nd IEEE Conference on Decision and Control*, December 2013.
- [23] C. Somarakis and J. S. Baras. Fixed point theory approach to exponential convergence in ltv continuous time consensus dynamics with delays. In *SIAM Control Conference*, July 2013.
- [24] C. Somarakis and J. S. Baras. A fixed point theory approach to multi-agent consensus dynamics with delays. Technical Report TR-2013-01, Univ. Of Md., College Park, USA, 2013.
- [25] C. Somarakis and J. S. Baras. On the convergence rate of non-linear consensus networks with delays. In *Proceedings of 9th IFAC Symposium on Nonlinear Control Systems*, September 2013.
- [26] C. Somarakis and J. S. Baras. LTI continuous time consensus dynamics with multiple and distributed delays. In *21st Mediterranean Conference on Control & Automation*, pages 1495–1501, June 2013.
- [27] J. Tsitsiklis, D. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, 31(9):803–812, 1986.
- [28] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75(6):1226–1229, 1995.