Scalability in Nonlinear Network Systems Affected by Delays and Disturbances

Shihao Xie ©, Giovanni Russo ©, Senior Member, IEEE, and Richard H. Middleton ©, Fellow, IEEE

Abstract—This article is concerned with the study of scalability in nonlinear heterogeneous networks affected by communication delays and disturbances. After formalizing the notion of scalability, we give two sufficient conditions to assess this property. Our results can be used to study leader–follower and leaderless networks and allow us to consider the case when the desired configuration of the system changes over time. We show how our conditions can be turned into design guidelines to guarantee scalability and illustrate their effectiveness via numerical examples.

Index Terms—Networks of autonomous agents, nonlinear systems, scalability of networks, stability.

I. INTRODUCTION

NETWORK systems have considerably evolved, increasing both their scale and the complexity of their topology [1]. Examples of large-scale networks include autonomous vehicles [2], robotic formations [3], and neural networks [4]. There is then no surprise that a large body of literature devoted to the study of collective behaviors has emerged, with, e.g., consensus and synchronization attracting much research attention [5]–[7].

In this context, a key challenge is the design of protocols that do not just guarantee stability (fulfillment of desired behavior for a fixed network) but also that the network is scalable. We use scalability to denote the preservation of desired stability properties (to be defined more formally in Section III) uniformly with respect to the size of the network. Scalability is then a fundamental requirement for network systems spanning from, e.g., platoons of vehicles to neural networks. For example, as noted in, e.g., [8] there is an intrinsic limit (this limit appears to be approximately 14% of the network size) for recurrent networks that precludes them to store an arbitrarily large number of memory patterns. Hence, increasing memory capacity in general requires increasing the number of neurons. However, if not properly designed, adding new layers/neurons might lead to the amplification of disturbances/biases as these propagate through the network [9]. Designing the network so that this is not just stable but also scalable avoids the onset of this undesirable behavior. Motivated by this, we introduce a notion of scalability for networks affected by delays, give sufficient conditions to assess this property, and show how our approach can be turned into design guidelines for scalability.

Manuscript received October 3, 2020; revised January 4, 2021; accepted January 18, 2021. Date of publication February 11, 2021; date of current version September 17, 2021. The work of Shihao Xie was supported in part by the Science Foundation Ireland (SFI) under Grant 16/IA/4610 and in part by SFI under Grant 13/RC/2094 (Lero). The work of Giovanni Russo was supported in part by SFI under Grant 13/RC/2094 (Lero). The work of Richard H. Middleton was supported by the Australian Research Council under Grant DP190102859. Recommended by Associate Editor G. Como. (Corresponding author: Giovanni Russo.)

Shihao Xie is with the School of Electrical and Electronic Engineering, University College Dublin, Dublin 4, Ireland (e-mail: shihao.xie1@ucdconnect.ie).
Giovanni Russo is with the Department of Information and Electric Engineering and Applied Mathematics, University of Salerno, 84084 Fisciano, Italy (e-mail: giovannorusso@unisa.it).
Richard H. Middleton is with the School of Electrical Engineering and Computing, University of Newcastle, Callaghan, NSW 2308, Australia (e-mail: Richard.Middleton@newcastle.edu.au).
Digital Object Identifier 10.1109/TCNS.2021.3058934

A. Related work

The study of how disturbances propagate within a network is a central topic for autonomous vehicle-following systems. In the context of platoons, much research effort has been devoted to the study of string stability [10]. The key idea behind the several definitions of string stability proposed in the literature [11] is that of giving upper bounds on certain deviations of the individual agents from a reference that are independent on the number of vehicles. See, e.g., [2] and [10]–[13] for recent results and a survey of the related literature. In the above works, results are obtained under the assumption that the network system is delay-free, and some extensions to strings affected by communication delays are given, e.g., [14] and [15] for homogeneous, disturbance-free, linear systems. Works on scalability for networks with arbitrary topologies are sparse when compared to works on string stability. For networks with linear agents, results include [1] and [16], where network coherence is characterized as a function of the number of its agents, [17] where performance deterioration in networks subject to external stochastic disturbances are considered, and [18] where certain network performance metrics are studied as a function of the number of edges. Other related works include these on leader-to-formation stability (see, e.g., [19]) that was introduced to characterize the behavior of (disturbance-free) formations with respect to the inputs provided by leaders and these on mesh stability (see, e.g., [20]) that offer a generalization of string stability to (linear and disturbance-free) networks with regular topologies. Other results include [21], where sufficient conditions for the scalability of delay-free leaderless networks with homogeneous agents interacting over regular topologies are introduced. Our proofs leverage contraction-theory arguments for time-delayed
systems. We recall [22], which shows, using the Euclidean metric, how contraction is preserved through certain time-delayed communications, and [23], which, by extending the integral quadratic constraints method, provides a tighter characterization of delays. Finally, we make use of max-separable metrics, and we refer the reader to, e.g., [24], where conditions for the synthesis of distributed controls are given by using separable metric structures.

B. Statement of the contributions

Our results are based on the key observation that, if a system is contracting with respect to a max-separable metric, then this is also scalable. We then give a set of sufficient conditions (independent on the bounds of the delays) for scalability of network systems consisting of heterogeneous nonlinear agents that communicate via possibly nonlinear protocols. The agents are affected by external disturbances and communication delays. Namely, our contributions can be summarized as follows.

1) For these delayed networks, we formalize the notions of $L_\infty$-scalable input-to-state stable ($L_\infty$-sISS) and $L_\infty$-scalable input-output stable ($L_\infty$-sIOS) networks.

2) We give two sufficient conditions to assess these properties. To the best of our knowledge, these are the first results that tackle the problem of guaranteeing scalability of nonlinear networks of heterogeneous agents affected by both disturbances and delays. Moreover, our results can be used to study both leader–follower and leaderless networks and also allow us to consider the case where the desired configuration of the system changes over time.

3) We show that our conditions can be turned into design guidelines for scalability. We do so by first showing how our approach can be used to design protocols that, while guaranteeing the tracking of a time-varying speed profile, ensure scalability for a network of mobile agents. We then show how the approach can be used to devise conditions on the activation functions (and their weights) to guarantee scalability of certain neural networks. Motivated by applications such as associative memory, where it is of interest to study stability of equilibrium when the network is forced by a constant input [25], the case we consider is when the recurrent network receives as input a constant vector, possibly affected by a bias, and the desired output is an equilibrium point (see also [26] and [27]). To the best of our knowledge, these are the first results that explicitly address scalability in neural networks.

II. MATHEMATICAL PRELIMINARIES

Let $A$ be an $m \times m$ real matrix. We denote by $\|A\|_p$ the matrix norm induced by the $p$-vector norm $\|\cdot\|_p$. We recall that (see, e.g., [28]) the matrix measure induced by $\|\cdot\|_p$ is defined as $\mu_p(A) := \max_{i \neq j} A_{ij}/\|A\|_p$.

The next result follows directly from [29]. We let $\| \cdot \|_S$ and $\mu_S(\cdot)$ be, respectively, any $p$-vector norm and its induced matrix measure on $\mathbb{R}^m$. In particular, the norm $\| \cdot \|_S$ is monotone, i.e., for any vector $x, y \in \mathbb{R}^n$, $x \leq y$ implies that $\|x\|_S \leq \|y\|_S$, where the inequality $x \leq y$ is componentwise.

**Lemma 1:** Consider the vector $\eta := [\eta_1^T, \ldots, \eta_N^T]^T$, $\eta_i \in \mathbb{R}^n$. We let $|\eta| := \|\eta_1\|_G, \ldots, \|\eta_N\|_G$, with $\|\cdot\|_G$ being norms on $\mathbb{R}^n$, and denote by $\|\cdot\|_G, \mu_G(\cdot)$ the matrix norm and measure induced by $\|\cdot\|_G$. Finally, let:

1) $A := (A_{ij})_{i,j=1}^N \in \mathbb{R}^{n \times n \times N}$, $A_{ij} \in \mathbb{R}^{n \times n}$; $\hat{A} := (\hat{A}_{ij})_{i,j=1}^N \in \mathbb{R}^{n \times n \times N}$, with $\hat{A}_{ij} := \mu_G(A_{ij})$ and $\hat{A}_{ij} := \|A_{ij}\|_{G_\ell,ij}$. $\|A_{ij}\|_{G_\ell,ij} := \sup_{|\eta_i| = 1} |A_{ij}| |\eta_i| |G_{ij}|$;

3) $A := (A_{ij})_{i,j=1}^N \in \mathbb{R}^{n \times n \times N}$, with $\hat{A}_{ij} := \|A_{ij}\|_{G_{\ell,ij}}$.

Then, $\mu_G(A) \leq \mu_S(\hat{A})$ and $\|A\|_G \leq \|\hat{A}\|_S$.

The next proposition follows from [30, Th. 2.4].

**Lemma 2:** Let $u := [-\tau_0, +\infty) \to \mathbb{R}^n$, $\tau_0 > +\infty$, and assume that $D^+ u(t) \leq a(u(t) + b \sup_{\tau(-\tau_0) \leq s \leq t} u(s) + c), t \geq 0$ with: 1) $\tau(t)$ being bounded and nonnegative, i.e., $0 \leq \tau(t) \leq \tau_0 \forall t$; 2) $u(t) = |\phi(t)| \forall t \in [-\tau_0, 0]$, where $\phi(t)$ is bounded in $[-\tau_0, 0]$; and 3) $a < 0 < b$, and $c \geq 0$. Assume that there exists some $\sigma > 0$ such that $a + b \leq -\sigma < 0 < c$. Then, $u(t) \leq \sup_{-\tau_0 \leq s \leq t} u(s) e^{-\lambda t} + \frac{c}{\sigma}$, where $\lambda := \inf_{t \geq 0} \{\lambda |\lambda(t) + a + be^{\lambda(t)}(t) = 0\}$ is positive.

III. STATEMENT OF THE CONTROL PROBLEM

We consider a network of $N$ heterogeneous agents, possibly receiving inputs from a set of $L$ leaders. The dynamics of the $i$th agent, $i = 1, \ldots, N$, is modeled via

\[ \dot{x}_i = f_i(x_i, t) + u_i(t) + b_i(x_i, t) d_i(t), \quad t \geq 0 \] (1)
\[ y_i = g_i(x_i) \] (2)

with $x_i \in \mathbb{R}^n$, $u_i(t) \in \mathbb{R}^n$ being the control protocol, and $d_i(t)$ being an $n$-dimensional signal modeling a (deterministic) external disturbance on the agent. In the dynamics, $b_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n \times n}$ models a time- and state-dependent disturbance intensity function. The function $f_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is the intrinsic dynamics of the agent, and $g_i : \mathbb{R}^n \to \mathbb{R}^m$ is the output function. The noise intensity is bounded, i.e., $\max_{i \leq i \leq N} \|b_i(x_i, t)\|_2 \leq b \forall t$ and $f_i(\cdot, \cdot)$, $g_i(\cdot)$ are smooth in their arguments. We
consider protocols of the form
\[
\begin{align*}
    u_i(t) &= \sum_{j \in N_i} h_{ij}^{(r)}(x_i(t), x_j(t), t) \\
    &+ \sum_{j \in N_i} h_{ij}(x_i(t), x_j(t), t) \\
    &+ \sum_{l \in L_i} h_{il}^{(s)}(x_i(t), x_l(t), t) \\
    &+ \sum_{l \in L_i} h_{li}(x_i(t), x_l(t), t)
\end{align*}
\]
(3)

with \(\tau(t) \leq \tau_0 \forall t, x_i(s) = \varphi_i(s), \) and \(\varphi_i(s) : \mathbb{R} \to \mathbb{R}^n\) being continuous and bounded \(\forall s \in [-\tau_0, 0] \forall i = 1, \ldots, N.\) In (3): (1) \(N_i\) denotes the set of neighbors of agent \(i\) and \(L_i\) is the set of leaders to which the \(i\)th agent is possibly connected; and (2) the functions \(h_{ij}^{(r)} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n\) and \(h_{ij}^{(s)} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n\) are the \textit{delay free} interagent and agent–leader coupling functions, respectively. Analogously, \(h_{ij}^{(r)} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n\) and \(h_{ij}^{(s)} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n\) are coupling functions for the delayed information. That is, in (3), we have separate couplings for the delay-free and delayed communication.

Remark 1: Situations where there is an overlap between a delay-free and delayed communication naturally arise in a number of applications. For example, in platooning [31], certain states (e.g., separation obtained by radar) from the neighbors might be available at an agent with a negligible/no delay. Other information, such as separation from more distant agents or neighbor control actions, may require communications and be subject to measurement/processing delays. Protocols of the form of (3) can also be used to study the popular neural network model of, e.g., [32]. In this case, agents are neurons and a delay-free communication coupling models an activation function from the closest neurons. The delayed communication can instead model interactions from neurons farther away in the network.

A. Control Goal

As in [13], we state our control goal in terms of a \textit{desired} solution of the unperturbed dynamics of (1)–(3). We let \(x^d_i(t) = [x^d_{i1}(t)^T, \ldots, x^d_{in}(t)^T]^T,\) with \(x^d_{ij}(t) = f_i(x^d_i(t), t),\) be the desired state/solution of the network when there are no disturbances and \(y_i(t) = [y^d_{i1}(t)^T, \ldots, y^d_{in}(t)^T]^T,\) with \(y^d_{ij}(t) = g_{ij}(x^d_{ij}(t)),\) be the desired output. Our goal is to design (3) so that (1) and (2) are either \(L_{\infty}\text{-sISS} or \(L_{\infty}\text{-sIOS}.

Definition 1: The closed-loop network system (1)–(3) is:

1) \(L_{\infty}\text{-sISS if there exists some class-\(\mathcal{K}\) function, \(\beta,\) and class-\(\mathcal{K}\) function, \(\gamma,\) such that, for all \(t \geq 0,\) \(\max_{i} |x_i(t) - x^d_i(t)|_2 \leq \beta(\max_{i} \sup_{-\tau_0 \leq s \leq 0} |x_i(s) - x^d_i(s)|_2, t) + \gamma(\sup_{-\tau_0 \leq s \leq 0} |x_i(s) - x^d_i(s)|_2, t)\) \(\forall N.\)
2) \(L_{\infty}\text{-sIOS if there exists some class-\(\mathcal{K}\) function, \(\kappa,\) and class-\(\mathcal{K}\) function, \(\gamma,\) such that, for all \(t \geq 0,\) \(\max_{i} |x_i(t) - x^d_i(t)|_2 \leq \kappa(\max_{i} \sup_{-\tau_0 \leq s \leq 0} |x_i(s) - x^d_i(s)|_2, t) + \gamma(\max_{i} |x_i(s) - x^d_i(s)|_2, t)\) \(\forall N.\)

Remark 2: We say that the network system is \(L_{\infty}\text{-sISS (or \(L_{\infty}\text{-sIOS)\) the above definition is fulfilled. The upper bounds in the definition give an estimate on the maximum deviation of state (output) from the desired configuration/solution. The functions \(\kappa,\) \(\beta,\) and \(\gamma\) are not dependent on the number of agents in the network, and such invariance of the bounds w.r.t. \(N\) is a key feature that differentiates scalability from the classic notions of input-to-state stability (ISS) and input–output stability (IOS). This guarantees that disturbances will not grow without bound as new agents are added, thus supporting the possibility of adding new agents.

IV. Technical Results

Our first result is a sufficient condition guaranteeing \(L_{\infty}\text{-sISS of (1)–(3).\) Whenever it is clear from the context, we omit the dependence of the state variables on the time.

Proposition 1: Consider network (1)–(3) with \(y_i(t) = x_i(t).\) Assume that \(\forall i = 1, \ldots, N\) and \(\forall \xi \geq 0,\) the following conditions are satisfied for some \(0 < \sigma < \sigma < c + \infty:\)

i) \(h_{ij}(x^d_i(t), x^d_j(t), t) = h_{ij}^{(r)}(x^d_i(t), x^d_j(t) - \tau(t), t), t = h_{il}^{(s)}(x^d_i(t), x_j(t), t) = h_{li}^{(s)}(x^d_i(t) - \tau(t), x_i(t), t), t)\)

ii) \(\mu_c(\partial_x f_i(x_i, t) + \sum_{j \in L_i} \partial_x h_{ij}(x_i, x_j, t) + \sum_{j \in N_i} \partial_x h_{ij}(x_i, x_j, t)) = -\sigma, \forall x_i, x_j \in \mathbb{R}^n.\)

iii) \(\sum_{j \in L_i} |\partial_x h_{ij}(x_i, x_j, t)|_2 + \sum_{j \in N_i} |\partial_x h_{ij}(x_i, x_j, t)|_2 \leq \sigma, \forall x_i, x_j \in \mathbb{R}^n.\)

Then, the system is \(L_{\infty}\text{-sISS. In particular, we have \(\forall \xi \geq 0\)}\)

\[
\max_i \left| x_i(t) - x^d_i(t) \right|_2 \leq \max \sup_{-\tau_0 \leq s \leq 0} \left| x_i(s) - x^d_i(s) \right|_2 e^{-\lambda t} + \frac{\tilde{b}}{\sigma - \tilde{b}} \left| x_i(t) - x^d_i(t) \right|_2 e^{-\lambda t}
\]

(4)

where \(0 < \tilde{\lambda} = \inf_{\xi \geq 0} \{ \lambda |y_i(t) - \sigma \leq y_i(t)(\tau(t)) = \lambda \}, \) \(x_i(t) = 0,\) \(x^d_i(t) = x^d_i(0),\) \(x_i(t) = \varphi_i(s), s \in [-\tau_0, 0], i = 1, \ldots, N.\)

In what follows, when we state the other scalability results, we omit that \(x(s) = \varphi_i(s) \forall s \in [-\tau_0, 0]\) and that \(x^d_i(t) = x^d_i(0) \forall s \in [-\tau_0, 0]\) as this is clear from the context. Before presenting the proof of the above result, we note the following.

Remark 3: Condition (i) implies that, at the desired solution, \(u_i(t) = 0.\) This rather common (see, e.g., [2], [3], and [5]) condition guarantees that the desired solution is a solution of the unperturbed dynamics. Conditions (ii) and (iii) give upper bounds on the matrix measure and matrix norm of the Jacobian of the controlled network system. These conditions imply that the Jacobian of the intrinsic dynamics and of the delay-free part of the protocol have a matrix measure that is \textit{sufficiently negative} to balance the presence of the delays. As we shall see, the conditions imply the existence of a norm in which the delay-free part of the dynamics has a negative matrix measure that is small enough to compensate the delays.

Remark 4: Interestingly, if the conditions of the proposition are satisfied, the scalability of the formation is guaranteed for any bounded delay. While scalability is guaranteed independently
the delay, the *convergence rate* \( \lambda \) depends on \( \tau(t) \). Also, we do not require any assumption on the differentiability of the delays. In this sense, our results relax an assumption made to study stability in, e.g., [33] and related references.

**Remark 5:** As in [21], the definition of scalability used in this article is independent on the topology of the interconnections between the agents. Also, in accordance to [21], designing a given network so that each agent has the number of neighbors independent of the network size can be leveraged to satisfy conditions (ii) and (iii) of Proposition 1. In general, adding connections between agents can lead to a loss of the scalability property (see also Fig. 3 and the related discussion in Section VI-A).

We are now ready to give the proof for Proposition 1.

**Proof:** We start with noting that, following condition (i), \( x_i^d(t) \) satisfies \( \dot{x}_i(t) - x_i^d(t) = f_i(x_i(t), t) \). Hence, we have

\[
\dot{x}_i(t) - x_i^d(t) = f_i(x_i(t), t) - f_i(x_i^d(t), t)
\]

\[+ \sum_{l \in \mathcal{E}_i} h_{il}(x_i(t), x_l(t), t) - \sum_{l \in \mathcal{E}_i} h_{il}(x_i^d(t), x_l(t), t)
\]

\[+ \sum_{l \in \mathcal{E}_i} h_{il}^\tau(x_i(t - \tau(t)), x_l(t - \tau(t)), t)
\]

\[- \sum_{l \in \mathcal{E}_i} h_{il}^\tau(x_i^d(t - \tau(t)), x_l(t - \tau(t)), t)
\]

\[+ \sum_{j \in N_i} h_{ij}(x_i(t), x_j(t), t) - \sum_{j \in N_i} h_{ij}(x_i^d(t), x_j^d(t), t)
\]

\[+ \sum_{j \in N_i} h_{ij}^\tau(x_i(t - \tau(t)), x_j(t - \tau(t)), t)
\]

\[- \sum_{j \in N_i} h_{ij}^\tau(x_i^d(t - \tau(t)), x_j^d(t - \tau(t)), t) + b_i(x_i(t), d_i(t))
\]

Let \( z(t) = [z_i^T(t), \ldots, z_N^T(t)]^T, \ z_i(t) = x_i(t) - x_i^d(t) \). Then, the dynamics of \( z(t) \) can be written as follows (see, e.g., [34]):

\[
\dot{z}(t) = A(t)z(t) + H(t)z(t - \tau(t)) + B(x,t)d(t)
\]

(5) where \( A(t) \) is an \( nN \times nN \) matrix consisting of the \( n \times n \) blocks defined \( \forall i,j = 1, \ldots, N \), as follows: 1) \( A_{ij}(t) := \int_0^t \partial f_i(e_i(\eta), t)d\eta + \sum_{l \in \mathcal{E}_i} \int_0^t \partial h_{il}(e_i(\eta), x_l(t), t)d\eta + \sum_{j \in N_i} \int_0^t \partial h_{ij}(e_i(\eta), e_j(\eta), t)d\eta \); and 2) \( A_{ij}(t) := \int_0^t \partial h_{ij}^\tau(e_i(\eta), e_j(\eta), t)d\eta \), where \( e_i(\eta) := \eta x_i + (1 - \eta) x_i^d \). Let us now study the error dynamics in (5). To this aim, we make use of Lemma 2 and define \( |z(t)|_G := ||z_1(t)||_2, \ldots, ||z_N(t)||_2 ||_\infty, \) which can be easily seen to be a vector norm. By taking the Dini derivative of \( |z(t)|_G \), from (5), we obtain

\[
D^+ |z(t)|_G := \limsup_{h \to 0^+} \frac{1}{h} (|z(t + h)| - |z(t)|)
\]

\[= \limsup_{h \to 0^+} \frac{1}{h} \left( |z(t) + H(t)z(t - \tau(t))| + hB(x,t)d(t)| - |z(t)| \right)
\]

\[\leq \mu_G(A)(|z(t)|_G) + \mu_G(H)(|z(t)|_G) \sup_{t - \tau(t) \leq s \leq t} |z(s)|_G
\]

\[+ \bar{b} \max_i \|d_i(\cdot)\|_\infty
\]

which was obtained by means of the triangle inequality and by using the fact that, from the definition of \( |z|_G \) and the boundedness of \( b_i(x_i(t), \|B(x,t)|_G \leq \max \sup_{x,t} \|b_i(x_i(t)|_2 \leq \bar{b}. In order to apply Lemma 2, we need to find the upper bound of \( \mu_G(A)(t) \) and \( \mu_G(H)(t) \). This can be computed via Lemma 1. Indeed, from such a result, it follows that \( \mu_G(A)(t) \leq \max \{\mu_2(A_{ii}(t)) + \sum_{j \in N_i} \|A_{ij}(t)|_2 \} \) and \( \mu_G(H)(t) \leq \max \{\sum_{j=1}^N \|H_{ij}(t)|_2 \} \). Moreover, conditions (ii) and (iii) imply that \( \max_i \{\mu_2(A_{ii}(t)) + \sum_{j \in N_i} \|A_{ij}(t)|_2 \} \leq -\sigma \) and \( \max_i \{\sum_{j=1}^N \|H_{ij}(t)|_2 \} \leq \sigma', \) for some \( 0 < \sigma' < \sigma < +\infty \). Hence, we get

\[
D^+ |z(t)|_G \leq -\sigma |z(t)|_G + \sup_{t - \tau(t) \leq s \leq t} |z(s)|_G
\]

\[+ \bar{b} \max_i \|d_i(\cdot)\|_\infty
\]

\[\forall N, \text{ which, by means of Lemma 2, yields}
\]

\[
|z(t)|_G \leq \sup_{-\tau_0 \leq s \leq 0} |z(s)|_G e^{-\sigma t} + \frac{\bar{b}}{\sigma - \sigma'} \max_i \|d_i(\cdot)\|_\infty
\]

\[\forall N, \text{ with } 0 < \hat{\lambda} = \inf_{s \geq 0} \{\lambda(t)| = -\sigma + \sigma e^{\lambda(t)} - \lambda(t) \}. \]

Since \( z(t) = x_i(t) - x_i^d(t) \), this completes the proof.

To further highlight a key difference between Definition 1 and the classical notion of ISS, consider the special case of a string of agents arranged in a cascading configuration. In this case, it is well known that the cascade of ISS agents is also ISS. However, this does not guarantee that the network is scalable. Indeed, Besselink and Johansson [35] give a counterexample (see Remark 3 therein), where perturbations grow without bound for a string of ISS systems when the number of agents increases. With the next result, we give a sufficient condition for \( L_\infty \)-sIOS.

**Proposition 2:** Consider the closed-loop network (1)–(3) and assume that conditions (i)–(iii) of Proposition 1 are satisfied and that, additionally, the functions \( g_i(\cdot) \) are Lipschitz with Lipschitz constant \( k_i \). Then, (1)–(3) are also \( L_\infty \)-sIOS.

**Proof:** The fulfillment of (i)–(iii) of Proposition 1 implies the upper bound in (4). Since the output functions are Lipschitz, we also have that \( |y_i - y_i^d|_2 := |g_i(x_i) - g_i(x_i^d)|_2 \leq k_i |x_i - x_i^d|_2 \). To obtain the desired bound, we need to show that

\[
|z(t)|_G \leq \sup_{-\tau_0 \leq s \leq 0} |z(s)|_G e^{-\sigma t} + \frac{\bar{b}}{\sigma - \sigma'} \max_i \|d_i(\cdot)\|_\infty
\]

\[\forall N, \text{ with } 0 < \hat{\lambda} = \inf_{s \geq 0} \{\lambda(t)| = -\sigma + \sigma e^{\lambda(t)} - \lambda(t) \}. \]

Since \( z(t) = x_i(t) - x_i^d(t) \), this completes the proof.

To further highlight a key difference between Definition 1 and the classical notion of ISS, consider the special case of a string of agents arranged in a cascading configuration. In this case, it is well known that the cascade of ISS agents is also ISS. However, this does not guarantee that the network is scalable. Indeed, Besselink and Johansson [35] give a counterexample (see Remark 3 therein), where perturbations grow without bound for a string of ISS systems when the number of agents increases. With the next result, we give a sufficient condition for \( L_\infty \)-sIOS.

**Proposition 2:** Consider the closed-loop network (1)–(3) and assume that conditions (i)–(iii) of Proposition 1 are satisfied and that, additionally, the functions \( g_i(\cdot) \) are Lipschitz with Lipschitz constant \( k_i \). Then, (1)–(3) are also \( L_\infty \)-sIOS.

**Proof:** The fulfillment of (i)–(iii) of Proposition 1 implies the upper bound in (4). Since the output functions are Lipschitz, we also have that \( |y_i - y_i^d|_2 := |g_i(x_i) - g_i(x_i^d)|_2 \leq k_i |x_i - x_i^d|_2 \). To obtain the desired bound, we need to show that

\[
|z(t)|_G \leq \sup_{-\tau_0 \leq s \leq 0} |z(s)|_G e^{-\sigma t} + \frac{\bar{b}}{\sigma - \sigma'} \max_i \|d_i(\cdot)\|_\infty
\]

\[\forall N, \text{ with } 0 < \hat{\lambda} = \inf_{s \geq 0} \{\lambda(t)| = -\sigma + \sigma e^{\lambda(t)} - \lambda(t) \}. \]

Since \( z(t) = x_i(t) - x_i^d(t) \), this completes the proof.
\[ x_i^d(t) = k \sup_{0 \leq t \leq T} x_i - x_i^d(0), \] where \( k := \max \{ k_i \} \). This, together with (4), immediately implies the result.

V. USING THE RESULTS TO DESIGN SCALABLE NETWORK SYSTEMS

We illustrate how our approach can be effectively used to design scalable networks. The first application we consider is concerned with the design of distributed protocols for a robotic formation. With the second application, we focus on ensuring that certain recurrent neural networks (RNNs) are scalable.

A. Formation Scalability

We now consider the problem of designing a control protocol guaranteeing scalability of the formation for a network of \( N \) mobile robots, while tracking a time-varying reference provided by a virtual leader. Each robot is modeled via a nonholonomic unicycle, and in particular, we adapt the popular model from [36] by embedding external disturbances

\[
\begin{align*}
\dot{p}_i^x &= v_i \cos \theta_i, \quad \dot{p}_i^y = v_i \sin \theta_i, \\
\dot{\theta}_i &= \omega_i \\
\dot{v}_i &= \frac{F_i + d_i^f}{m_i}, \quad \dot{\omega}_i = \frac{Q_i + d_i^\omega}{I_i}.
\end{align*}
\]

(6)

In the model, \( p_i(t) := [p_i^x(t), p_i^y(t)]^T \) is the inertial position, \( v_i(t) \) is the linear speed, \( \theta_i(t) \) is the heading angle, \( \omega_i(t) \) is the angular velocity, \( m_i \) is the mass, \( F_i(t) \) is the applied force input, \( Q_i(t) \) is the applied torque input, \( I_i(t) \) is the moment of inertia, \( d_i^f(t) \) is the external force disturbance, and \( d_i^\omega(t) \) is the external torque disturbance. In (6), the disturbance \( d_i^f(t) \) models uncertainties due to, e.g., unmodeled friction forces, while \( d_i^\omega(t) \) models external disturbances due to, e.g., wind. We aim at controlling the hand position of the \( i \)-th robot. As shown in the Appendix, the dynamics (6) can be feedback linearized, yielding

\[
\begin{align*}
\dot{\chi}_i &= A_i \chi_i + \nu_i(t) + b_i(t)d_i(t), \\
\eta_i &= C_i \chi_i
\end{align*}
\]

(7)

with control \( \nu_i(t) := [\begin{smallmatrix} \nu_i^x \\ \nu_i^y \\ \nu_i^\theta \\ \nu_i^\omega \end{smallmatrix}] \), \( \chi_i := [\chi_i^1, \chi_i^2, \chi_i^3, \chi_i^4]^T = [p_i^x + l_i \cos \theta_i, p_i^y + l_i \sin \theta_i, v_i \cos \theta_i - l_i w_i \sin \theta_i, v_i \sin \theta_i + l_i w_i \cos \theta_i]^T 
\]

\[
A_i := \begin{bmatrix}
0_{2 \times 2} & I_2 \\
0_{2 \times 2} & 0_{2 \times 2}
\end{bmatrix}, \
\]

\[
b_i(t) := \begin{bmatrix} I_2 \\ 0_{2 \times 2} \end{bmatrix}, \quad d_i(t) := \begin{bmatrix} \dot{\chi}_i \\ \nu_i(t) \end{bmatrix}, \quad C_i := [I_2 \ 0_{2 \times 2}]
\]

(8)

\[ d_i(t) = [0_{2 \times 2} \ I_2 \ 0_{2 \times 2} \ I_2 \ 0_{2 \times 2} \ 0_{2 \times 2} \ 0_{2 \times 2} \ 0_{2 \times 2}]. \]

In what follows, we design the control \( \nu_i(t) \) so that the above network system is \( L_\infty \)-sIOS, and hence, it has a scalable formation. We consider the popular class of protocols considered in, e.g., [37], where communications between robots are affected by delay and robots have access to a reference trajectory (i.e., hand position and speed) provided by a virtual leader, \( \chi(t) \). In particular, \( \eta_i(t) := [\chi_i^1, \chi_i^2]^T \) denotes the hand position of the leader at time \( t \) and \( \nu_i(t) := [\chi_i^3, \chi_i^4]^T \) is the corresponding smooth speed signal. Within the formation, the hand position of the \( i \)-th robot needs to keep some desired offset from the neighbors and from \( \eta_i(t) \), while, at the same time, tracking the acceleration and speed provided by the virtual leader. That is, the desired solution for the \( i \)-th robot within the formation when there are no perturbations/delays, i.e., \( \chi^d_i(t) := [\chi_{i,1}^d, \chi_{i,2}^d, \chi_{i,3}^d, \chi_{i,4}^d]^T \), is such that: 1) the desired hand position of the robot \( \eta^d_i(t) := [\chi_{i,1}^d, \chi_{i,2}^d]^T \) keeps the desired offsets; 2) the corresponding speed is \( \nu^d_i(t) = [\chi_{i,3}, \chi_{i,4}]^T \); and 3) it satisfies

\[
\chi_{i}^d = \begin{bmatrix} 0_{2 \times 2} & I_2 \\
0_{2 \times 2} & 0_{2 \times 2}
\end{bmatrix} \chi_{i}^d + \begin{bmatrix} 0_{2 \times 1} \\
0_{2 \times 1}
\end{bmatrix} \nu^d_i(t).
\]

(9)

1) Protocol Design: We consider control protocols of the form

\[
\begin{align*}
\dot{\nu}_i(t) &= \dot{\nu}_i(t) + \sum_{j \in N_i} h_{ij} \left( \chi_i(t) - \tau(t) \right), \chi_j(t) - \tau(t)), t \\
&+ \bar{h}_d \left( \chi_i(t), \chi(t), t \right)
\end{align*}
\]

(10)

where the coupling functions \( h_{ij} \) : \( R^4 \times R^4 \times R^4 \rightarrow R^2 \) and \( \bar{h}_d : R^4 \times R^4 \times R^4 \rightarrow R^2 \) can be nonlinear and are smooth and where all the agents are affected by the same delay as in [33]. In what follows, we let \( h_{ij} \left( \chi_i(t) - \tau(t), \chi_j(t) - \tau(t) \right), t := \begin{bmatrix} h_{ij}(\chi_i(t) - \tau(t), \chi_j(t) - \tau(t), t) \end{bmatrix}^T \
\bar{h}_d(\chi_i(t), \chi(t), t) \end{bmatrix}^T \) and \( \bar{h}_d(\chi_i(t), \chi(t), t) \end{bmatrix} = \begin{bmatrix} h_{ij}(\chi_i(t) - \tau(t), \chi_j(t) - \tau(t), t) \end{bmatrix}^T \). Then, with the following results, we establish a sufficient condition for \( L_\infty \)-sIOS of the closed-loop dynamics (7)–(10). The results are stated in terms of the following matrix:

\[
T_i = \begin{bmatrix} I_2 \\
0_{2 \times 2} \end{bmatrix}, \quad \alpha_i > 0.
\]

(11)

Proposition 3: Consider the network of mobile agents (7), (8) controlled by (10). Assume that the coupling functions \( h_{ij} \), \( \bar{h}_d \) satisfy (10). Then, the following conditions for some \( 0 < \sigma < \sigma < +\infty \) and some vector \([\alpha_1, \ldots, \alpha_N] \) of nonnegative constants:

\[
C_1 \quad h_{ij} \left( \chi_i^d(t - \tau(t)), \chi_j^d(t - \tau(t)), t \right) = \bar{h}_d(\chi_i(t), t) = 0;
\]

\[
C_2 \quad \mu_{ij} T_i A_i T_i^{-1} + T_i \partial h_d(\chi_i(t), \chi(t), T_i^{-1}) \leq -\sigma, \forall \chi_i, \chi_i \in R^4;
\]

\[
C_3 \quad \sum_{j \in N_i} T_i \partial h_d(\chi_i(t), \chi_j(t), T_i^{-1}) \leq -\sigma, \forall \chi_i, \chi_i \in R^4.
\]

Then, the network is \( L_\infty \)-sIOS and in particular \( \forall t \geq 0:
\]

\[
\max_{i} \left| \eta_i(t) - \eta_i^d(t) \right|_2 \leq K \max_{i} \sup_{-\tau \leq \tau \leq 0} \left| \chi_i(s) - \chi_i^d(s) \right|_2 e^{-\lambda t} + K \max_{i} \max_{\sigma_i(t)} \left| \partial h_d(\chi_i(t), \chi(t), t) \right| \leq \sigma \] \]

where

\[
K := \max_{i} \left\{ \frac{\sigma_{\min}(T_i)}{\min_{i} \left\{ \sigma_{\min}(T_i) \right\}} \right\},
\]

\[
\max_{i} \sup_{\tau \in [-\tau, 0]} \left| \partial h_d(\chi_i(t), \chi(t), t) \right| \leq \sigma \]

\[
\lambda := \inf_{\tau \geq 0} \{ \lambda | \lambda(t) - \bar{\sigma} + \sigma e^{\lambda(t)} \alpha_i(t) | = 0 \}.
\]

Proof: We prove the result via Proposition 2, and again, we omit the explicit dependence of the state variables on time as this is clear from the context. Clearly, the output function is
Lipschitz, and therefore, in order to apply the result, we only need to show that the conditions of Proposition 1 are satisfied. First, we let \( \chi_i(t) := [\chi_{i1}(t), \ldots, \chi_{Ni}(t)]^T \) be the desired solution of the network, corresponding to the desired formation. We then note that, by means of C1, \( \chi_i(t) \) is a solution of (9). Also, the fulfillment of C1 implies the fulfillment of condition (i) in Proposition 1. In order to continue with the proof, for the dynamics (7), (8), we consider \( \forall i = 1, \ldots, N \), the coordinate transformation \( \tilde{\chi}_i(t) := T_i \chi_i(t) \) with \( T_i \) defined in (11). In particular, by writing (7) and (8) in these new coordinates, one can note that C2 and C3 are equivalent to (ii) and (iii) of Proposition 1. Therefore, we have

\[
\max_i \| \tilde{\chi}_i(t) - \tilde{\chi}_i^d(t) \|_2 \leq \max_i \sup_{-\tau_0 \leq s \leq 0} \| \tilde{\chi}_i(s) - \tilde{\chi}_i^d(s) \|_2 e^{-\lambda t} + \max_i \{ \sigma_{\max}(T_i) \} \frac{\max_i \| d_i(\cdot) \|_\infty}{\lambda - \sigma} \forall N, \text{ with } \tilde{\chi}_i^d(t) \text{ being the desired solution of robot } i \text{ in the new coordinates.}
\]

To obtain the upper bound, we used the fact that:
1) \( \| TB(t) \| \leq \| T \| \| B(t) \| \leq \max_i \{ \sigma_{\max}(T_i) \} \| B \| \) and \( B(t) \) are the \( N \times N \) block diagonal matrices having on their main diagonal the blocks \( T_i \) and \( b_i(t) \), respectively;
2) \( \max_i |\tilde{\chi}_i(t) - \tilde{\chi}_i^d(t)| < \frac{\lambda}{\max_i \{ \sigma_{\max}(T_i) \}} \forall N, \text{ with } \tilde{\chi}_i^d(t) \text{ being the desired solution of robot } i \text{ in the new coordinates.}
\]

\[
\max_i \| \tilde{\chi}_i(t) - \tilde{\chi}_i^d(t) \|_2 \leq \frac{\lambda}{\max_i \{ \sigma_{\max}(T_i) \}} \sup_{-\tau_0 \leq s \leq 0} \| \tilde{\chi}_i(s) - \tilde{\chi}_i^d(s) \|_2 e^{-\lambda t} + \max_i \{ \sigma_{\max}(T_i) \} \frac{\max_i \| d_i(\cdot) \|_\infty}{\lambda - \sigma} \forall N
\]

proving the result as \( |\eta_i(t) - \eta_i^d(t)| \leq \| \chi_i(t) - \chi_i^d(t) \|_2 \).

\textbf{B. Scalability in Cohen–Grossberg RNNs}

We now consider the problem of designing scalable neural networks. In particular, we focus on Cohen–Grossberg neural network, which are widely used for, e.g., pattern recognition, associative memories [32], and have Hopfield neural networks as a special case. The model we consider is:

\[
\begin{align*}
\dot{x}_i &= p_i(x_i(t))(-c_i(x_i(t)) + \sum_{j=1}^{N} a_{ij} g_j(x_j(t)) + \sum_{j=1}^{N} b_{ij} \tilde{g}_j^{(\tau)}(x_j(t - \tau(t))) + u_i + d_i(t)) \\
&= p_i(x_i(t))(-c_i(x_i(t)) + \sum_{j=1}^{N} a_{ij} g_j(x_j(t)) + \sum_{j=1}^{N} b_{ij} \tilde{g}_j^{(\tau)}(x_j(t - \tau(t))) + u_i + d_i(t))
\end{align*}
\]

where \( x_i(t) \in \mathbb{R} \) is the state of the \( i \)th neuron; \( \tau(t) \leq \tau_0 \) is the time-varying transmission delay associated with information transmission; \( p_i(\cdot) \) is the amplification function, which is assumed to be positive and bounded with \( p \leq p_i(x_i(t)) \leq \overline{p} p_i(\cdot) \); \( a_{ij} \in \mathbb{R} \) and \( b_{ij} \in \mathbb{R} \) are the corresponding neuron connection weights; \( u_i \) is the (possible) constant input to the neuron; \( d_i(t) \) is the exogenous disturbance. The disturbance can model environmental perturbations, adversarial attacks, and data biases. In what follows, \( x^* := [x_1^*, \ldots, x_N^*]^T \) is the desired network equilibrium point for a given input \( u = [u_1, \ldots, u_N]^T \).

\textbf{Proposition 4}: Consider the RNN (12). Assume that there exist some \( \sigma, \overline{\rho} > 0 \) such that \( \forall i = 1, \ldots, N \) and \( \forall t \geq 0 \):

\[
\begin{align*}
C1 - c_i(x_i(t)) + \sum_{j=1}^{N} a_{ij} g_j(x_j(t)) + \sum_{j=1}^{N} b_{ij} \tilde{g}_j^{(\tau)}(x_j(t)) + u_i = 0; \\
C2 - \partial_t c_i(x_i(t)) + a_{i1} \partial_t g_1(x_1(t)) + \sum_{j \neq i} |a_{ij} \partial_t g_j(x_j(t))| \leq -\sigma \\
\forall x_i, x_j \in \mathbb{R}; \\
C3 \sum_{j=1}^{N} |b_{ij} \tilde{g}_j^{(\tau)}(x_j(t))| \leq \sigma \forall x_j \in \mathbb{R}; \\
C4 \overline{\rho} \overline{p} \leq \rho^2.
\end{align*}
\]

Then, the network is \( \mathcal{L}_\infty \)-sISS. In particular \( \forall t \geq 0 \):

\[
\max_i \| x_i(t) - x_i^* \| \leq \max_i \sup_{-\tau_0 \leq s \leq 0} \| x_i(s) - x_i^* \| e^{-\lambda t} + \frac{\underline{\rho} \overline{p}}{\overline{\rho} \overline{p} - \rho^2} \max_i \| d_i(\cdot) \|_\infty \forall N
\]

where \( 0 < \lambda = \inf_{t \geq 0} \{ x(t) \lambda(t) - t \rho^2 - \overline{\rho} \overline{p} \sigma^{(\tau)}(t) \} = 0 \).

\textbf{Proof}: We start with noticing that condition C1, together with the fact that the \( p_i(\cdot) \) are bounded, implies that \( \dot{x}_i^*(t) = 0 = p_i(x_i^*(t))(-c_i(x_i^*(t)) + \sum_{j=1}^{N} a_{ij} g_j(x_j^*(t)) + \sum_{j=1}^{N} b_{ij} \tilde{g}_j^{(\tau)}(x_j^*(t - \tau(t))) + u_i) = p_i(x_i(t))(-c_i(x_i^*) + \sum_{j=1}^{N} a_{ij} g_j(x_j^*) + \sum_{j=1}^{N} b_{ij} \tilde{g}_j^{(\tau)}(x_j^*) + u_i) \), with \( x_i(t) \) being a solution of (12). Hence, we have

\[
\dot{x}_i(t) - \dot{x}_i^*(t) = p_i(x_i(t))\left(-c_i(x_i(t)) + c_i(x_i^*) + \sum_{j=1}^{N} a_{ij} g_j(x_j(t)) - \sum_{j=1}^{N} a_{ij} g_j(x_j^*) + \sum_{j=1}^{N} b_{ij} \tilde{g}_j^{(\tau)}(x_j(t - \tau(t))) - \sum_{j=1}^{N} b_{ij} \tilde{g}_j^{(\tau)}(x_j^*) + d_i(t) - d_i^*(t)\right)
\]

Again, we let \( z(t) = [z_1^T(t), \ldots, z_N^T(t)]^T \), \( z_i(t) = x_i(t) - x_i^* \). Then, the dynamics for \( z(t) \) can be written as \( \dot{z}(t) = P(x)A(t)z(t) + H(t)(z(t) - \tau(t)) + d(t) \), where \( P(x) \) is the diagonal matrix having \( p_i(x_i(t)) \) on its main diagonal and where \( A(t) \) is an \( N \times N \) matrix having entries defined \( \forall i, j = 1, \ldots, N \) as:
1) \( A_{ii}(t) := \int_0^t \partial_t c_i(x_i(t) + (1 - \eta)z_i(t))dn + \int_0^t \partial_t \tilde{g}_i(x_i(t) + (1 - \eta)z_i(t))dn \); and
2) \( A_{ij}(t) := \int_0^t \partial_t \tilde{g}_i(x_i(t) + (1 - \eta)z_i(t))dn \).

Also, \( d(t) = [d_1^T(t), \ldots, d_N^T(t)]^T \) and \( H(t) = (H_{ij})_{i,j=1}^N \) is the \( N \times N \) matrix with the elements:
1) \( H_{ii}(t) := \int_0^t \tilde{g}_i^{(\tau)}(x_i(t) + (1 - \eta)z_i(t))dn \); and
2) \( H_{ij}(t) := \int_0^t \tilde{g}_i^{(\tau)}(x_i(t) + (1 - \eta)z_i(t))dn \).

We then consider the Dini derivative of \( |z(t)|_\infty \), and this yields

\[
D^+ \| z(t) \|_\infty := \limsup_{h \rightarrow 0^+} \frac{1}{h} \{ |z(t + h)| - |z(t)| \} = \limsup_{h \rightarrow 0^+} \frac{1}{h} \{ |z(t) + hP(x)A(t)z(t) + hP(x)H(t)(z(t) - \tau(t)) + d(t)|_\infty - |z(t)|_\infty \}
\]
\[ \limsup_{t \to \infty} \frac{1}{t} \left( \|I + H(t)\|_\infty - 1 \right) |z(t)|_\infty \]
\[ + \bar{p} \|H(t)\|_\infty \sup_{t - \tau(t) \leq s \leq t} |z(s)|_\infty + \bar{p} \sup_{t} |d(t)|_\infty \]
\[ = \mu_\infty \left( P(x)A(t) \right) |z(t)|_\infty \]
\[ + \bar{p} \|H(t)\|_\infty \sup_{t - \tau(t) \leq s \leq t} |z(s)|_\infty + \bar{p} \sup_{t} |d(t)|_\infty \]

where we used the fact that \( \|P(x)\|_\infty = \max_i |p_i(x_i)| \leq \bar{p} \).

Moreover, by definition of \( \mu_\infty \), we obtain the following:
\[ \mu_\infty \left( P(x)A(t) \right) = \max_i \left\{ \frac{1}{\lambda_i} \left( \sum_{j \neq i} \lambda_j a_{ij} \right) + p_i(x_i) \right\} \]
\[ \leq \max_i \left\{ \frac{1}{\lambda_i} \left( \sum_{j \neq i} \lambda_j a_{ij} \right) + p_i(x_i) \right\} \]
\[ \leq \max_i \left\{ \frac{1}{\lambda_i} \left( \sum_{j \neq i} \lambda_j a_{ij} \right) + p_i(x_i) \right\} \]
\[ \leq \max_i \left\{ \frac{1}{\lambda_i} \left( \sum_{j \neq i} \lambda_j a_{ij} \right) + p_i(x_i) \right\} \]

Finally, \( C4 \) makes it possible to apply Lemma 2, and this implies the desired upper bound.

1) Hopfield Neural Networks: Hopfield neural networks are a special case of (12) when \( p_i(x_i) = 1 \) and \( c_i(x_i(t)) = c_i x_i(t) \) \( \forall i \).

The resulting model is then
\[ \dot{x}_i = -c_i x_i(t) + \sum_{j=1}^N a_{ij} g_j(x_j(t)) \]
\[ + \sum_{j=1}^N b_{ij} g_j^{(\gamma)}(x_j(t)) + u_i + d_i(t) \]

(14)

where \( i = 1, \ldots, N \). We let again \( x^e := [x_1^e, \ldots, x_N^e]^T \) be the desired equilibrium and give the following.

**Corollary 1:** Consider the Hopfield RNN (14). Assume that there exist some \( \bar{\sigma}, \sigma > 0 \) such that \( \forall i = 1, \ldots, N \) and \( \forall t \geq 0 \):
\[ C1 \quad -c_i x_i^e + \sum_{j=1}^N a_{ij} g_j(x_j^e) + \sum_{j=1}^N b_{ij} g_j^{(\gamma)}(x_j^e) + u_i \leq 0 \]
\[ C2 \quad -c_i + a_{ii} \partial_t g_i(x_i) + \sum_{j \neq i} |a_{ij} \partial_t g_j(x_j)| \leq -\bar{\sigma}, \forall x_i, x_j \in \mathbb{R} \]
\[ C3 \quad \sum_{j=1}^N |b_{ij} \partial_t g_j^{(\gamma)}(x_j)| \leq \sigma, \forall x_j \in \mathbb{R} \]
\[ C4 \quad \bar{\sigma} > \sigma \]

Then, the network is \( L_\infty \)-sISSL. In particular \( \forall t \geq 0 \):
\[ \max_i |x_i(t) - x_i^e| \leq \max_i \sup_{-\tau(t) \leq s \leq 0} |x_i(s) - x_i^e| e^{-\lambda t} \]
\[ + \frac{1}{\bar{\sigma} - \sigma} \max_i |d_i(t)|_\infty \]

where \( 0 < \lambda = \inf_{t \geq 0} \{ \lambda \lambda(t) - \bar{\sigma} + \sigma \partial \lambda(t)(t) \} = 0 \).

**Proof:** It follows directly from Proposition 4.

**Remark 6:** Essentially, Corollary 1 states that, if there are disturbances on the network, and the matrix measure induced by \( | \cdot |_\infty \) is considered to study stability, then not only the system is stable, as shown in [38], but it is also scalable.

**VI. NUMERICAL EXAMPLES**

Code, data, and parameters are available online.\(^1\)

\(^1\)https://github.com/GIOVRUSO/Control-Group-Code
With these parameters, we first investigate how the maximum deviation of the hand position (from its desired position) changes as a function of the number of agents in the formation. To this aim, starting with a formation of one circle, we repeatedly simulated the formation by increasing, at each simulation, the number of circles. Then, for each simulation, we recorded the maximum deviation experienced on each circle and finally plot the maximum deviation on each circle across all the simulations.

The result of this process is illustrated in the left panel of Fig. 2. Such a figure clearly shows that the disturbance on the first circle is not amplified across the other circles, in accordance with our theoretical predictions. As next step, we also investigate how, for the formation, the maximum deviation of the hand position changes as a function of the delay $\tau$. The results are illustrated in the right panel of Fig. 2, which shows that deviations stay bounded when the delay increases (the figure is obtained for the formation with six circles).

Finally, with our second set of simulations, we further investigate our scalability conditions by considering again the formation in Fig. 1 this time with 14 circles and 1) each robot on the ith circle connected to the one ahead and behind on the same circle and with the closest robot on circle $i-1$ (if any); 2) one robot on the inner circle (i.e., circle 1) is affected by the same disturbance used in Fig. 2; and 3) the delay set to $\tau = 0.1$ s. We simulated the network first with the same control weights used in Fig. 2 (i.e., satisfying the conditions of Proposition 3) and then with a set of control weights that make the network stable but not scalable. As clearly shown in the top panel of Fig. 3, scalability prohibits the amplification of perturbations propagating through the network. Instead, when the network is designed to be stable but not scalable, the disturbances grow when propagating before being attenuated (see the middle panel of Fig. 3). Finally, the unstable behavior illustrated in the bottom panel of Fig. 3 has been obtained by considering the same protocol and control weights used in the top panel of the figure but this time with each agent connected to all the others.

Interestingly, the simulations in the bottom panel of Fig. 3 show that, perhaps counterintuitively, a scalable network can be made unstable if new connections are added. We also note that the unstable behavior numerically observed in the bottom panel of Fig. 3 cannot be explained with our sufficient conditions for scalability. A related phenomenon has also been recently observed in the context of synchronization of diffusively coupled delay-free networks. In particular, in [39] and [40], it has been shown how a stable synchronization manifold can be made unstable by increasing the connections between certain nodes. With respect to this, an approach to theoretically explain the behavior observed in the bottom panel of Fig. 3 might be that of extending the sufficient conditions for desynchronization of [39] and [40].

**B. Designing Scalable Hopfield Neural Networks**

We now turn our attention to the problem of designing an $L_\infty$-sISS scalable Hopfield neural network. We consider (14) with 60 neurons and 1) each neuron connected to all the others; 2) $c_i = 10 \forall i$; and 3) nonnegative weights. All the activation functions are affected by the delay $\tau(t) = \tau = 1s$ (i.e., $a_{ij} = 0 \forall i,j$, in the model). Also, all the neurons have $g(\tau(x) = \tanh(x)$ as the activation function. In order to numerically validate the conditions of Corollary 1, we first computed a set of weights verifying conditions C2–C4 of Corollary 1. Then, we simulated the network without any disturbance and with nonnegative inputs $u_i$ (the specific inputs are at the repository). This was done to find the unique stable equilibrium, toward which the network converges (this is indeed the desired equilibrium; note how C1 is intrinsically satisfied by such equilibrium point). The behavior of the network when there are no disturbances is illustrated in the top panel of Fig. 4. The network behavior when the network is affected by disturbances is instead shown in the bottom panel of Fig. 4. In Fig. 4, the deviations of $x_i(t)$ s with respect to the equilibrium of the unperturbed network are shown. In Fig. 4,

---

1 We leave the study of this open problem for our future research.
2 Weights available at https://github.com/GIOVRUSSO/Control-Group-Code
55 neurons were perturbed at time 5 and 15 s, with constant disturbances having a random amplitude between 0 and 10 and duration of 1 s. Again, the figure illustrates that disturbances are attenuated within the network, in accordance with the findings. Finally, we also considered the Hopfield network shown in the top panel of Fig. 5. The activation functions were hyperbolic tangents, and these were all affected by delays. We set \( \tau(t) = \tau = 0.1 \) s, \( a_{ij} = 0 \), and \( b_{ij} = 15 \) for the connections shown in the top panel of Fig. 5 and first picked \( c_i = 27 \) \( \forall i \) so that the network was stable but not scalable. Then, we picked \( c_i = 32 \), which allowed to fulfill the conditions of Corollary 1, hence making the network scalable. As clearly illustrated in the bottom panels of Fig. 5, in accordance with our results, when the network is affected by disturbances, scalability prohibits their amplification within the network.

VII. CONCLUSION AND DISCUSSION

We considered networks of possibly nonlinear heterogeneous agents coupled via possibly nonlinear protocols affected by delays and disturbances. For these networks, after introducing the notions of \( L_\infty \)-sISS and \( L_\infty \)-sSOS, we presented two sufficient conditions to assess these properties. The conditions can be turned into design guidelines, and we used our results to: 1) design distributed control protocols able to guarantee both tracking of a time-varying reference and \( L_\infty \)-sSOS; and 2) design the activation functions (and their weights) of certain RNNs so that these are \( L_\infty \)-sISS. The effectiveness of the results was illustrated via simulations. Besides considering heterogeneous delays, future work might involve: 1) devising scalability conditions that take into account bounds on the delays and investigate the conservativeness of our conditions; 2) investigating, inspired by certain recent literature (see, e.g., [9] and [41]) on stable RNNs to model nonlinear input–output sequences, scalability for complex, time-varying, input–output patterns; and 3) stochastic disturbances. Finally, motivated by the numerical findings reported in the bottom panel of Fig. 3, our research also includes devising sufficient conditions for the loss of scalability.

APPENDIX

The derivations are inspired from [36], where the same model was considered but without disturbances. We let \( x_i = [p_i^x, p_i^\theta, v_i^\theta, \omega_i]^T, u_i = [F_i, Q_i]^T, \) and \( d_i = [d_i^F, d_i^Q]^T \) and aim at controlling the hand position \( \eta_i = p_i + l_i [\cos \theta_i, \sin \theta_i]^T, \) where \( l_i \) is the distance of the hand position from the inertial position. We differentiate \( \eta_i \) twice to obtain

\[
\ddot{\eta}_i = \begin{bmatrix}
-v_i \omega_i \sin \theta_i - l_i \omega_i^2 \cos \theta_i \\
v_i \omega_i \cos \theta_i - l_i \omega_i^2 \sin \theta_i
\end{bmatrix} + \frac{1}{m_i} \begin{bmatrix}
\cos \theta_i - \frac{1}{\omega_i} \sin \theta_i
\end{bmatrix} \begin{bmatrix}
F_i \\
Q_i
\end{bmatrix},
\]

Considering the diffeomorphism \( \chi_i = T_i(x_i) = [p_i^x + l_i \cos \theta_i, p_i^\theta + l_i \sin \theta_i, v_i \cos \theta_i - l_i \omega_i \sin \theta_i, v_i \sin \theta_i + l \omega_i \cos \theta_i, \theta_i]^T = [\chi_{i,1}, \chi_{i,2}, \chi_{i,3}, \chi_{i,4}, \chi_{i,5}]^T, \) we obtain \( \dot{\chi}_{i,1} = \chi_{i,3}, \dot{\chi}_{i,2} = \chi_{i,4}, \) and

\[
\begin{bmatrix}
\dot{\chi}_{i,3} \\
\dot{\chi}_{i,4} \\
\dot{\chi}_{i,5}
\end{bmatrix} = \begin{bmatrix}
-v_i \omega_i \sin \chi_{i,5} - l_i \omega_i^2 \cos \chi_{i,5} \\
v_i \omega_i \cos \chi_{i,5} - l_i \omega_i^2 \sin \chi_{i,5}
\end{bmatrix} + \frac{1}{m_i} \begin{bmatrix}
\cos \chi_{i,5} - \frac{1}{\omega_i} \sin \chi_{i,5}
\end{bmatrix} \begin{bmatrix}
\bar{u}_i \\
\bar{\bar{d}}_i
\end{bmatrix},
\]

\[
\dot{\chi}_{i,5} = -\frac{1}{2l_i} \chi_{i,3} \sin \chi_{i,5} + \frac{1}{2l_i} \chi_{i,4} \cos \chi_{i,5}
\]
with \( \eta_i = [\chi_{i,1}, \chi_{i,2}]^T \). The feedback linearizing control is

\[
\bar{u}_i = \left( \frac{1}{m_i} \cos \chi_{i,5} - \frac{I_i}{I} \sin \chi_{i,5} \right)^{-1} \times \left( \bar{v}_i(t) - \left[ -v_i \omega_i \sin \chi_{i,5} - \frac{1}{2} \omega_i^2 \cos \chi_{i,5} \right] \right).
\]

Hence, noticing that the zero dynamics is stable yields the following reduced dynamics:

\[
\begin{pmatrix}
\dot{x}_{i,1} \\
\dot{x}_{i,2} \\
\dot{x}_{i,3} \\
\dot{x}_{i,4}
\end{pmatrix} = \bar{v}_i(t) + \bar{b}_i(t) d_i(t)
\]

with \( \eta_i = [\chi_{i,1}, \chi_{i,2}]^T \) and where

\[
\bar{b}_i(t) := \left[ \frac{1}{m_i} \cos \theta_i(t) - \frac{I_i}{I} \sin \theta_i(t) \right].
\]

This is the dynamics considered in Section V-A, where \( \chi_i := [\chi_{i,1}, \chi_{i,2}, \chi_{i,3}, \chi_{i,4}] \).

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers and the Associate Editor for their constructive feedbacks.

REFERENCES

[1] E. Tegling, P. Mitra, H. Sandberg, and B. Bamieh, “On fundamental limitations of dynamic feedback control in regular large-scale networks,” \( \text{IEEE Trans. Autom. Control} \), vol. 64, no. 12, pp. 4936–4951, Dec. 2019.

[2] S. Stüdli, M. Seron, and R. Middleton, “From vehicular platoons to general networked systems: String stability and related concepts,” \( \text{Ann. Rev. Control} \), vol. 44, pp. 157–172, 2016.

[3] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, “Flocking in fixed and switching networks,” \( \text{IEEE Trans. Autom. Control} \), vol. 52, no. 5, pp. 863–868, May 2007.

[4] T. Shen and I. R. Petersen, “Linear threshold discrete-time recurrent neural networks: Stability and globally attractive sets,” \( \text{IEEE Trans. Autom. Control} \), vol. 50, no. 9, pp. 1520–1533, Sep. 2004.

[5] R. Olfati-Saber and R. M. Murray, “Consensus problems in networks of agents with switching topology and time-delays,” \( \text{IEEE Trans. Autom. Control} \), vol. 51, no. 4, pp. 783–795, Apr. 2006.

[6] R. Wang and I. R. Manchester, “Robust contraction analysis of nonlinear systems via differential IQC,” in \( \text{Proc. IEEE 58th Conf. Decis. Control} \), 2019, pp. 6766–6771.

[7] H. S. Shiro moto, M. Revay, and I. R. Manchester, “Distributed nonlinear control design using separable control contraction metrics,” \( \text{IEEE Trans. Control Netw. Syst.} \), vol. 6, no. 4, pp. 1281–1290, Dec. 2019.

[8] C. Zhou, X. Zeng, J. Yu, and H. Jiang, “A unified associative memory model based on external inputs of continuous recurrent neural networks,” \( \text{Neurocomputing} \), vol. 186, pp. 44–53, 2016.

[9] L. Wen, Y. Yu, and W. Wang, “Generalized Halanay inequalities for dissipativity of volterra functional differential equations,” \( \text{J. Math. Anal. Appl.} \), vol. 347, pp. 169–178, 2008.

[10] B. Besselink and K. H. Johansson, “String stability and a delay-based spacing policy for vehicle platoons subject to disturbances,” \( \text{IEEE Trans. Control Netw. Syst.} \), vol. 5, no. 1, pp. 197–211, Apr. 2014.

[11] D. A. Burbano-L. G. Russo, and M. di Bernardo, “Pinning controllability of complex network systems with noise,” \( \text{IEEE Trans. Control Netw. Syst.} \), vol. 6, no. 2, pp. 874–883, Jun. 2019.

[12] V. Follì, M. Leonetti, and G. Ruocco, “On the maximum storage capacity of the Hopfield model,” \( \text{Front. Comp. Neurosci.} \), vol. 10, 2017, Art. no. 144.

[13] F. Bonassi, E. Terzi, M. Farina, and R. Scattolini, “LSTM neural networks: Input to state stability and probabilistic safety verification,” \( \text{Proc. Mach. Learn. Res.} \), vol. 120, pp. 85–94, 2020.

[14] D. Swaroop and J. K. Hedrick, “String stability of interconnected systems,” \( \text{IEEE Trans. Autom. Control} \), vol. 41, no. 3, pp. 349–357, Mar. 1996.

[15] S. Feng, Y. Zhang, S. E. Li, Z. Cao, H. X. Liu, and L. Li, “String stability for vehicular platoon control: Definitions and analysis methods,” \( \text{Ann. Rev. Control} \), vol. 47, pp. 81–97, 2019.

[16] S. Knorn, A. Donaire, J. C. Agüero, and R. H. Middleton, “Passivity-based control for multi-vehicle systems subject to string constraints,” \( \text{Automatica} \), vol. 50, pp. 3224–3230, 2014.
[38] J. Cao and Y. Wan, “Matrix measure strategies for stability and synchronization of inertial BAM neural network with time delays,” Neural Netw., vol. 53, pp. 165–172, 2014.

[39] G. Russo, “Loss of coordination in complex directed networks: An incremental approach based on matrix measures,” Int. J. Robust Nonlinear Control, vol. 28, no. 1, pp. 120–131, 2018.

[40] G. Russo, “How to desynchronize quorum-sensing networks,” Phys. Rev. E, vol. 95, no. 4, Apr. 2017, Art. no. 042312.

[41] M. Revay and I. R. Manchester, “Contracting implicit recurrent neural networks: Stable models with improved trainability,” Proc. Mach. Learn. Res., vol. 120, pp. 393–403, 2020.

Shihao Xie received the B.Eng. degree in aircraft power engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2017, and the M.Sc. degree in autonomous vehicle dynamics and control from Cranfield University, Cranfield, U.K., in 2018. He is currently working toward the Ph.D. degree in electrical and electronic engineering with University College Dublin, Dublin, Ireland. His research interests include distributed control of networked systems with its application to automotive engineering and neural networks.

Giovanni Russo (Senior Member, IEEE) received the Ph.D. degree in control theory from the University of Naples Federico II, Naples, Italy, in 2010. He then joined Ansaldo STS as a System Engineer. From 2012 to 2015, he was the Lead System Engineer and Integrator of the Honolulu Rail Transit Project, USA. From 2015 to 2018, he was at IBM Research Ireland as a Research Staff Member in Optimization, Control and Decision Science. From 2018 to 2020, he was a Lecturer in Cyber-Physical Systems with University College Dublin, Dublin, Ireland. He is currently an Associate Professor with the University of Salerno, Italy. His main research interests include the intersection of control and information science with main applications involving intelligent transportation and cyber-physical systems.

Dr. Russo is a member of the Board of Editors of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS and an Associate Editor for the IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS.

Richard H. Middleton (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Newcastle, Callaghan NSW, Australia, in 1987. He was a Research Professor at the Hamilton Institute, The National University of Ireland, Maynooth, Ireland, from 2007 to 2011 and the Head of the School of Electrical Engineering and Computing, University of Newcastle, from 2015 to 2020. He is currently the Director of the ARC Training Centre for Food and Beverage Supply Chain Optimisation. His research interests include a broad range of control systems theory and applications, including communications systems, control of distributed systems, optimization, and systems biology.

Dr. Middleton was the President of the IEEE Control Systems Society (CSS) in 2011. He has served as the Program Chair (IEEE Conference on Decision and Control (CDC), 2006), the Co-General Chair (CDC 2017) CSS Vice-President Membership Activities, and Vice-President Conference Activities. He is a Fellow of the International Federation of Automatic Control.