Automatica 86 (2017) 205–211

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

On the stability and convergence of a class of consensus systems with a nonlinear input[☆]

Mingming Liu [a], Fabian Wirth [b], Martin Corless [c], Robert Shorten [a]

[a] School of Electrical, Electronic, and Communications Engineering, University College Dublin, Dublin, Ireland
[b] The Faculty of Computer Science and Mathematics, University of Passau, Passau, Germany
[c] School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN 47907-2045, USA

ARTICLE INFO

Article history:
Received 18 February 2016
Received in revised form 18 May 2017
Accepted 21 July 2017
Available online 28 September 2017

Keywords:
Nonlinear systems
Optimisation
Convergence proofs

ABSTRACT

We consider a class of consensus systems driven by a nonlinear input. Such systems arise in a class of Internet of Things (IOT) applications. Our objective in this paper is to determine conditions under which a certain partially distributed system converges to a Lur'e-like scalar system, and to provide a rigorous proof of its stability. Conditions are derived for the non-uniform convergence and stability of such a system and an example is given of a speed advisory system where such a system arises.

1. Introduction

We consider nonlinear discrete-time systems described by

$$x(k+1) = P(k)x(k) + \mu\big(r - g(x(k))\big)e \tag{1}$$

where $k = 0, 1, 2, \ldots$, $x(k) \in \mathbb{R}^n$, $P(k)$ is a $n \times n$ row stochastic matrix, $e = [1\ 1\ \ldots\ 1]^T$, μ and r are scalars while g is a scalar valued function. Eq. (1) arises in consensus problems subject to an output constraint. It basically says that if consensus is achieved, it must be achieved subject to the equilibrium constraint $g(x^*) = r$. Eq. (1) is of interest as it arises in many situations in the study of the *internet of things* (IOT). For example, in some situations a group of agents are asked to achieve a fair allocation of a constrained resource; TCP is an algorithm that strives to achieve this objective in internet congestion control. A second application arises when one wishes to optimise an objective function subject to certain privacy constraints. Sometimes, for reasons of privacy, one does not attempt to solve such problems in a fully distributed manner. Neither, for reasons of robustness, scale, and communication overhead, does one attempt to solve them in a centralised manner. Rather, one uses a mix of local communication, and limited broadcast information, to solve these problems in a manner that conceals the private information of each of the individual agents. Implicit and explicit consensus algorithms that exploit local and global communication strategies are proposed and studied in Knorn, Corless, and Shorten (2011) and Knorn, Stanojevic, Corless, and Shorten (2009). Eq. (1) is perhaps the simplest algorithm of the explicit consensus algorithm with inputs, admitting a very simple intuitive understanding. It is well known that a row stochastic matrix P operates on a vector $x \in \mathbb{R}^n$ such that $\max(x) - \min(x) \geq \max(Px) - \min(Px)$ where $\max(x)$ and $\min(x)$ are defined as the maximum and minimum component in vector x, respectively. Since the addition of $\big(r - g(x(k))\big)e$ does not affect this contraction, intuition suggests that $x_i(k) - x_j(k) \to 0$ as k increases and eventually, the dynamics of (1) will be governed by the following scalar **Lur'e system**:

$$y(k+1) = y(k) + \mu(r - g(y(k)e)), \tag{2}$$

with $x_i(k) = y(k)$ asymptotically for all i. Intuition further suggests that, as long as (2) is stable, then so is (1). Arguments along these lines, in support of (1), are given in Knorn et al. (2009) and Knorn et al. (2011) . However, these arguments are not complete in the sense that certain important properties are assumed to hold true. Our objective therefore in this brief note is to establish conditions on the function g for which global uniform asymptotic stability is assured, and to rigorously prove the resulting assertions.

1.1. Comments on related literature

Before proceeding it is prudent to discuss related work.

(i) *Cascade Systems :* The setup we study can be viewed as a more general case of the systems studied in Lu, Atay, and Jost (2007).

☆ This work was in part supported by Science Foundation Ireland under grant 11/PI/1177. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Michael M. Zavlanos under the direction of Editor Christos G. Cassandras.

E-mail addresses: mingming.liu@ucd.ie (M. Liu), fabian.lastname@uni-passau.de (F. Wirth), corless@purdue.edu (M. Corless), robert.shorten@ucd.ie (R. Shorten).

http://dx.doi.org/10.1016/j.automatica.2017.08.003
0005-1098/

In contrast to the assumptions of that paper we do not require differentiability of the system maps and we state conditions which ensure global convergence. Further related work concerns stability analysis of nonlinear cascades (Loría & Nešić, 2002; Loría & Nešić, 2003; Nešić & Loría, 2004); we will comment on the relation to this literature in the system description and when stating results.

(ii) *Consensus :* The setup we discuss is obviously connected to work on consensus. While the literature on consensus is too rich to give a complete survey here, recent surveys are available in Mesbahi and Egerstedt (2010) and Olfati-Saber, Fax, and Murray (2007). Here we briefly note that the standard problem studied in this literature involves conditions that guarantee convergence of solutions to the consensus subspace. In this paper, we study the more specific problem of convergence to a specific point in which consensus is reached. Standard results ensuring consensus will therefore not apply in any classical sense to the problem studied here. Furthermore, a standard assumption in the case of a consensus system is a connectivity assumption on the communication graph. For a discussion of conditions used in this area we refer the reader to Blondel, Hendrickx, Olshevsky, and Tsitsiklis (2005), Mesbahi and Egerstedt (2010) and Moreau (2005).

(iii) *Optimisation :* Note, the idea to use consensus techniques to obtain approximations of optimal solutions has already been studied in other papers; see Burger, Notarstefano, and Allgower (2014), Nedic, Ozdaglar, and Parrilo (2010) and Shi, Johansson, and Hong (2013) for some work in this direction. For example, in Burger et al. (2014), a constraint exchange based consensus algorithm was proposed, where the authors combined the ideas with dual decomposition and cutting-plane methods to solve convex and robust distributed optimisation problems via polyhedral approximations. While this approach is applicable to more general problems than those studied here, it is also more complex. Further, we note that many of the other techniques in the literature rely on the use of individual constraint sets for the agents and projections onto that set. These reduce to standard consensus in the case that no constraints are present.

(iv) *Convergence :* As a further difference we point out that we give convergence results which can be non-uniform, whereas the authors in Nedic et al. (2010) point out that they rely on uniform convergence.

2. Notation, conventions and preliminary results

2.1. Notation

We denote the standard basis in $\mathbb{R}^n$ by the vectors $e_1, \ldots, e_n$. Note that $e = \sum_{i=1}^n e_i$. A matrix $P \in \mathbb{R}^{n\times n}$ is called row stochastic, if all its entries are nonnegative and if all its row sums equal one. The row sum condition is equivalent to $Pe = e$, that is, e is an eigenvector of P corresponding to the eigenvalue 1. Hence there is a single transformation which achieves upper block triangularisation of all row stochastic matrices. Let $\{v_2, \ldots, v_n\}$ be a basis for the $n-1$ dimensional subspace $e^\perp := \{x \in \mathbb{R}^n : e^T x = 0\}$. Then $\{e, v_2, \ldots, v_n\}$ is a basis of $\mathbb{R}^n$. Consider now the transformation matrix $T := \begin{bmatrix} e & v_2 & \ldots & v_n \end{bmatrix}$ which represents a change of basis from the standard basis to the new basis. Under this transformation, a row stochastic matrix P is transformed as follows:

$$T^{-1}PT = \begin{bmatrix} 1 & c \\ 0 & Q \end{bmatrix}. \tag{3}$$

2.2. Facts about consensus

Given a sequence of row stochastic matrices $\{P(k)\}_{k\in\mathbb{N}}$, consider the time-varying linear system

$$x(k+1) = P(k)x(k). \tag{4}$$

A solution of (4) is represented by the left products of the matrix sequence in the following sense: a sequence $\{x(k)\}_{k\in\mathbb{N}}$ is a solution of (4) corresponding to the initial condition $x(0) = x_0$ if and only if for all $k \in \mathbb{N}$,

$$x(k) = \Phi(k)x_0 \tag{5}$$

where

$$\Phi(k) := P(k-1)\cdots P(0) \qquad \text{for all} \quad k \in \mathbb{N}. \tag{6}$$

The sequence $\{P(k)\}_{k\in\mathbb{N}}$ is called weakly ergodic if the difference between each pair of rows converges to zero, i.e. if for all i, j we have

$$\lim_{k\to\infty} \left(e_j^T - e_i^T\right) \Phi(k) = 0. \tag{7}$$

This is equivalent to system (4) being a consensus system, that is, every solution $\{x(k)\}_{k\in\mathbb{N}}$ of (4) satisfies

$$\lim_{k\to\infty} x_j(k) - x_i(k) = 0 \tag{8}$$

for all i, j. The sequence $\{P(k)\}_{k\in\mathbb{N}}$ is strongly ergodic if it is weakly ergodic and, in addition, the limit $\lim_{k\to\infty}\Phi(k)$ exists; we denote this limit by Φ_∞. Due to a result of Chatterjee and Seneta (1977), weak and strong ergodicity are equivalent for left products of row stochastic matrices. This is equivalent to every solution of (4) satisfying

$$\lim_{k\to\infty} x(k) \in E, \tag{9}$$

or, equivalently, $\lim_{k\to\infty}\Phi(k)x_0 \in E$ for all $x_0 \in \mathbb{R}^n$ where $E := \text{span}\{e\}$ is the space of consensus vectors. We call the sequence $\{P(k)\}_{k\in\mathbb{N}}$ strongly ergodic for all initial times, if all tail sequences $\{P(k)\}_{k=k_0}^\infty$ are strongly ergodic for all $k_0 \in \mathbb{N}$. Note that a sequence can be strongly ergodic and not strongly ergodic for all initial times. For instance, if one of the matrices in the sequence has rank 1 and all the subsequent matrices are the identity matrix. Using the transformation (3) a system equivalent to (4) is given by

$$\begin{aligned} z(k+1) &= T^{-1}P(k)Tz(k) \\ T^{-1}P(k)T &:= \begin{bmatrix} 1 & c(k) \\ 0 & Q(k) \end{bmatrix}. \end{aligned} \tag{10}$$

It is then clear that $\{P(k)\}$ is strongly ergodic if and only if

$$\lim_{k\to\infty} Q(k)\ldots Q(0) = 0. \tag{11}$$

A useful property in the study of products of row stochastic matrices is the observation that for any row stochastic matrix P we have

$$\min(x) \le \min(Px) \le \max(Px) \le \max(x) \tag{12}$$

for all $x \in \mathbb{R}^n$, where for any vector $y \in \mathbb{R}^n$,

$$\min(y) := \min\{y_1, \ldots, y_n\}, \ \max(y) := \max\{y_1, \ldots, y_n\}.$$

Introducing the function

$$V(x) := \max(x) - \min(x), \tag{13}$$

(12) implies that $V(Px) \le V(x)$. Also, the sequence $\{P(k)\}_{k\in\mathbb{N}}$ is strongly ergodic, if and only if

$$\lim_{k\to\infty} V\big(\Phi(k)x_0\big) = 0 \tag{14}$$

for all $x_0 \in \mathbb{R}^n$ where $\Phi(k)$ is given by (6). Note that any vector $x \in \mathbb{R}^n$ can be uniquely decomposed as

$$x = \bar{x}e + x_\perp, \tag{15}$$

where

$$\bar{x} := (1/n)e^T x \tag{16}$$

is the mean of the components of x and

$$x_\perp := x - \bar{x}e\,. \tag{17}$$

Note that $\bar{x}e \in E$ and $e^T x_\perp = 0$; hence

$$\operatorname{dist}(x, E) = \|x_\perp\| \tag{18}$$

where $\operatorname{dist}(x, E) := \inf\{\|x - z\| \ : z \in E\}$ is the distance of a vector $x \in \mathbb{R}^n$ to the consensus set E and $\|x\| = \sqrt{x^T x}$. Note also that

$$V(x) = V(x_\perp)$$

and for any vector $z \in \mathbb{R}^n$ and any row-stochastic matrix $P \in \mathbb{R}^{n\times n}$ we have

$$\|z\|_\infty = \max_{1\le i\le n} |z_i| \quad \text{and} \quad \|P\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |P_{ij}| = 1$$

where P_{ij} is the entry in the ith row and jth column of the matrix P. As the following results rely on the existence of strongly ergodic sequences, it is of course of interest to have criteria for the occurrence of such a sequence. This is discussed extensively in the literature and here we can only discuss a limited number of references. A number of such criteria can be found in Chatterjee and Seneta (1977). Relations of this notion to ergodicity or the Dobrushin coefficient are discussed in Ipsen and Selee (2011) and the references given therein. In addition, in the consensus literature there are numerous results on the convergence of iterated averaging, see e.g. Blondel et al. (2005, Theorem 1), and the survey given in Olfati-Saber et al. (2007, Section III).

3. Consensus under feedback

Consider a sequence of row stochastic matrices $\{P(k)\}_{k\in\mathbb{N}}$ and a continuous function $G : \mathbb{R}^n \to \mathbb{R}$. Then the system,

$$\begin{aligned} x(k+1) &= F(k, x(k)) \\ F(k, x) &:= P(k)x + G(x)e \end{aligned} \tag{19}$$

can be regarded as a consensus system under feedback. In later statements, further differentiability assumptions will be imposed on G as required. If we apply the similarity transformation defined in (3), then in the new coordinates, $y \in \mathbb{R}, z \in \mathbb{R}^{n-1}$, given by $x = T[y\ z^T]^T$, we obtain

$$\begin{aligned} &y(k+1) = y(k) + G(T[y(k)\,z(k)^T]^T) + c(k)z(k) \\ &z(k+1) = Q(k)z(k)\,. \end{aligned} \tag{20}$$

The class of cascaded systems studied in Loría and Nešić (2002), Loría and Nešić (2003) and Nešić and Loría (2004) encompasses this system formulation. We note that in all these references and in subsequent literature based on these papers (Lee & Jiang, 2006), it is assumed that the subsystems are globally uniformly asymptotically stable. We do not require this assumption. Associated with (19) we consider the one-dimensional system

$$\begin{aligned} y(k+1) &= h(y(k)) \\ h(y) &:= y + G(ye)\,, \end{aligned} \tag{21}$$

which is seen to be the one dimensional subsystem in the cascade (20) corresponding to $z(k) = 0$. This is the aforementioned Lur'e system and, as we shall see, the dynamics of the consensus system (19) is strongly related to the dynamics of (21). Unless stated otherwise we consider the systems (19) and (21) with initial time $k_0 = 0$. Comments on results that hold uniformly with respect to all initial times are made where appropriate.

3.1. Local stability results

Lemma 1. *Let $\{P(k)\}_{k\in\mathbb{N}}$ be a sequence of row stochastic matrices and $G : \mathbb{R}^n \to \mathbb{R}$. If $\{y(k)\}_{k\in\mathbb{N}}$ is a solution of (21) then $\{y(k)e\}_{k\in\mathbb{N}}$ is a solution of (19).*

Proof. This follows from $P(k)e = e$. □

The next result tells us that the consensus system under feedback (19) is also a consensus system. We omit the straightforward proof of this observation.

Lemma 2. *If $\{P(k)\}_{k\in\mathbb{N}}$ is a strongly ergodic sequence of row stochastic matrices, then for every solution $\{x(k)\}_{k\in\mathbb{N}}$ of (19) we have*

$$\lim_{k\to\infty} \operatorname{dist}(x(k), E) = 0\,. \tag{22}$$

We now consider the local stability of (19) and see that it is determined by the stability of the induced system (21) on the consensus space. As we have no global concerns no Lipschitz property of G is required. Initially, it is sufficient that G be continuous.

Theorem 1. *Let $\{P(k)\}_{k\in\mathbb{N}}$ be a strongly ergodic sequence of row stochastic matrices and $G : \mathbb{R}^n \to \mathbb{R}$ be continuous. Suppose that y^* is a locally asymptotically stable fixed point of the one dimensional system (21). Then y^*e is a locally asymptotically stable fixed point at time $k_0 = 0$ for (19). Furthermore, if the sequence $\{P(k)\}_{k\in\mathbb{N}}$ is strongly ergodic for all initial times, then y^*e is asymptotically stable for all initial times $k_0 \in \mathbb{N}$.*

Proof. Suppose that y^* is a locally asymptotically stable fixed point for system (21). Let W be a local Lyapunov function which guarantees this stability property. That is, $W(y^*) = 0$ and there is an open neighbourhood U of y^* such that $W(y) > 0$ and $W(h(y)) < W(y)$ for all $y \in U \setminus \{y^*\}$. Without loss of generality we may assume U to be a forward invariant set of (21), i.e., if $y \in U$ then $h(y) \in U$. For $\varepsilon > 0$ such that $W^{-1}([0, \varepsilon]) \subset U$ is a compact set we may choose $\delta > 0$ sufficiently small, so that

$$W(h(y) + d) < \varepsilon \qquad \text{for} \qquad W(y) < \epsilon \quad \text{and} \quad |d| \le \delta\,.$$

This is possible by continuity of all the functions involved and by the decay property of the Lyapunov function W. We note that for any $x \in \mathbb{R}^n$, $Px = P(\bar{x}e + x_\perp) = \bar{x}e + Px_\perp$, where $\bar{x}$ was defined as the mean of the entries of $x \in \mathbb{R}^n$ (recall from (16)). Hence

$$\overline{Px} - \bar{x} = \bar{x} + \overline{Px_\perp} - \bar{x} = (1/n)e^T P x_\perp\,. \tag{23}$$

Given a sufficiently small $\varepsilon > 0$ and an appropriate δ as above, choosing $\eta > 0$ such that $V(x) \le \eta$ and $W(\bar{x}) \le \epsilon$ implies that for any row stochastic matrix P

$$|\overline{Px} - \bar{x}| + |G(x) - G(\bar{x}e)| < \delta\,.$$

This is possible by the estimate for

$$\|Px - x\|_\infty = \|Px_\perp - x_\perp\|_\infty \le \|Px_\perp\|_\infty + \|x_\perp\|_\infty \le 2\|x_\perp\|_\infty$$

and by uniform continuity of G on a bounded neighbourhood of y^*e. Consider now the neighbourhood of y^*e given by

$$N_\varepsilon := \{x \in \mathbb{R}^n \ : \ \bar{x} \in U,\ W(\bar{x}) < \varepsilon, V(x) < \eta\}\,.$$

We claim that N_ε is forward invariant at all times $k \in \mathbb{N}$. Indeed, if $x(k) \in N_\varepsilon$, then we obtain

$$\begin{aligned} \bar{x}(k+1) &= \overline{P(k)x(k)} + G(x(k)) \\ &= \bar{x}(k) + G(\bar{x}(k)e) + d \\ &= h(\bar{x}(k)) + d \end{aligned}$$

where $d = \overline{P(k)x(k)} - \bar{x}(k) + G(x(k)) - G(\bar{x}(k)e)$. Hence $|d| < \delta$ from which it follows that $W(\bar{x}(k+1)) < \varepsilon$. Referring to the argument in the proof of Lemma 2

$$V(x(k+1)) = V(P(k)x(k)) \le V(x(k)) < \eta.$$

As ε, η were arbitrary, this shows stability of y^*e. To show local attractivity, let $x_0 \in N_\varepsilon$ for $\varepsilon > 0$ sufficiently small so that stability holds. Note that by Lemma 2 and by stability we have that $\omega(x_0) \subset Ue \subset E$ where $\omega(x_0)$ is the ω-limit set of the solution corresponding to x_0 and $Ue := \{ye : y \in U\}$ is a subset of E. Suppose that $ye \in \omega(x_0)$ and $y \neq y^*$. Then as the trajectory starting in ye converges to y^*e it follows that $y^*e \in \omega(x_0)$. However, the assumption that y^*e and ye are in the ω-limit set contradicts the stability of y^*e. Hence $\{x(k)\}_{k\in\mathbb{N}}$ converges to y^*e. □

We now extend the previous result to local exponential stability. To this end we call a sequence of row stochastic matrices $\{P(k)\}_{k\in\mathbb{N}}$ exponentially ergodic if it is strongly ergodic and there exist scalars $M \ge 1$, $0 < r < 1$ such that for all $k \in \mathbb{N}$

$$\|\Phi(k) - \Phi_\infty\| \le M r^k .$$

The sequence is called uniformly exponentially ergodic, if it is strongly ergodic for all initial times and there exist constants M, r so that for all $k_0 \in \mathbb{N}$ there exists a matrix Φ_∞ so that for all $k > k_0$ we have $\|\tilde{\Phi}(k, k_0) - \Phi_\infty\| \le M r^{(k-k_0)}$; where $\tilde{\Phi}(k, k_0) := P(k-1)\cdots P(k_0)$.

Theorem 2. *Let $\{P(k)\}_{k\in\mathbb{N}}$ be an exponentially ergodic sequence of row stochastic matrices and $G : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable. Suppose that y^* is a locally exponentially stable fixed point of the one dimensional system (21). Then, y^*e is a locally exponentially stable fixed point at time $k_0 = 0$ for (19). If the sequence $\{P(k)\}_{k\in\mathbb{N}}$ is uniformly exponentially ergodic, then y^*e is locally uniformly exponentially stable.*

Proof. Consider the linearisation of the one-dimensional map defining (21). By the assumption of exponential stability it must satisfy

$$|h'(y^*)| < 1 \tag{24}$$

where $h'(y^*) = 1 + DG(y^*e)e$ and DG is the derivative of G, which we interpret as a row vector. We now compute the derivative of F with respect to x at $x = y^*e$ and time k to obtain

$$\frac{\partial F}{\partial x}(k, y^*e) = P(k) + eDG(y^*e). \tag{25}$$

If we now consider the transformation T which results in (10) and using $T^{-1}e = e_1$ we see that

$$T^{-1}\frac{\partial F}{\partial x}(k, y^*e)T = \begin{bmatrix} 1 & c(k) \\ 0 & Q(k) \end{bmatrix} + e_1 DG(y^*e)T . \tag{26}$$

Two things are noticeable when considering this equation. First the resulting transformed matrix is of the form

$$\begin{bmatrix} \lambda & \tilde{c}(k) \\ 0 & Q(k) \end{bmatrix} , \tag{27}$$

where only the first row is affected by G and λ is independent of k. Secondly,

$$\lambda = 1 + DG(y^*e)e = h'(y^*) . \tag{28}$$

Hence $|\lambda| < 1$. By assumption $\|Q(k)Q(k-1)\dots Q(0)\| \le Mr^k$ for suitable constants $M \ge 1$ and $r \in (0, 1)$. It now follows that the linearised system of (19) at the fixed point y^*e is exponentially stable. It follows by standard linearisation theory that the nonlinear system is locally exponentially stable at y^*e. If the sequence $Q(k)Q(k-1)\dots Q(0)$ converges to zero uniformly exponentially, this shows local uniform exponential stability of y^*e for the nonlinear system. □

3.2. Global stability results

To obtain global stability results we first need the following boundedness result.

Lemma 3. *Let $\{P(k)\}_{k\in\mathbb{N}}$ be a strongly ergodic sequence of row stochastic matrices and suppose that $G : \mathbb{R}^n \to \mathbb{R}$ is continuous and satisfies the following conditions.*

(i) *There exists an $\varepsilon > 0$ such that G satisfies a Lipschitz condition with constant $L > 0$ on the set*

$$B_\varepsilon(E) := \{x \in \mathbb{R}^n \ : \ \mathrm{dist}(x, E) \le \varepsilon\} .$$

(ii) *There exist constants $\beta, \gamma > 0$ such that*

$$|h(y)| \le |y| - \gamma \qquad \text{when} \quad |y| \ge \beta$$

where $h(y) = y + G(ye)$.

Then every trajectory of (19) is bounded.

Proof. Consider any solution $\{x(k)\}_{k\in\mathbb{N}}$ of (19) with $x(0) = x_0$. By Lemma 2 there exists a $k_0 \in \mathbb{N}$ such that $x(k) \in B_\varepsilon(E)$ for all $k \ge k_0$. We can express $x(k)$ as $x(k) = \bar{x}(k)e + x_\perp(k)$ where $\bar{x}(k) = (1/n)e^T x(k)$ and $x_\perp(k) := x(k) - \bar{x}(k)e$. It follows from (22) that $\lim_{k\to\infty} \|x_\perp(k)\| = 0$. Hence boundedness of the sequence $\{\bar{x}(k)\}_{k\in\mathbb{N}}$ implies boundedness of $\{x(k)\}_{k\in\mathbb{N}}$. Considering the evolution of $\bar{x}(k)$ we obtain that, for $k \ge k_0$,

$$\begin{aligned}
|\bar{x}(k+1)| &= |\overline{P(k)x(k)} + G\big(\bar{x}(k)e + x_\perp(k)\big)| \\
&\le |\bar{x}(k) + G(\bar{x}(k)e)| + |\overline{P(k)x(k)} - \bar{x}(k)| \\
&\quad + |G\big(\bar{x}(k)e + x_\perp(k)\big) - G(\bar{x}(k)e)| \\
&\le |h(\bar{x}(k))| + |(1/n)e^T P(k)x_\perp(k)| + L\|x_\perp(k)\| \\
&\le |h(\bar{x}(k))| + \tilde{L}\|x_\perp(k)\|
\end{aligned}$$

where $l := \sup_{k\in\mathbb{N}} (1/n)\|e^T P(k)\|$ and $\tilde{L} := l + L$. Hence

$$|\bar{x}(k+1)| \le |h(\bar{x}(k))| + \tilde{L}\|x_\perp(k)\| .$$

It now follows from hypothesis (ii) that whenever $|\bar{x}(k)| \ge \beta$, we must have

$$|\bar{x}(k+1)| \le |\bar{x}(k)| - \gamma + \tilde{L}\|x_\perp(k)\|.$$

Since $\lim_{k\to\infty} \|x_\perp(k)\| = 0$, there exists a $k_* \ge k_0$ such that $\tilde{L}\|x_\perp(k)\| \le \gamma$ for all $k > k_*$. Thus,

$$|\bar{x}(k+1)| \le |\bar{x}(k)| \qquad \text{when} \quad k \ge k_* \text{ and } |\bar{x}(k)| \ge \beta.$$

This implies boundedness of $\{\bar{x}(k)\}_{k\in\mathbb{N}}$ and completes the proof. □

Remark 1. As an example of a general class of functions which satisfy hypothesis (ii) of Lemma 3, consider any strict contraction mapping h on $\mathbb{R}$, i.e., for a suitable constant $c \in (0, 1)$,

$$|h(x) - h(y)| \le c|x - y| , \quad \forall x, y \in \mathbb{R}.$$

By the Banach contraction theorem, there is a unique fixed point y^* such that $h(y^*) = y^*$. Hence,

$$\begin{aligned}
|h(y)| &\le |h(y) - y^*| + |y^*| \le c|y - y^*| + |y^*| \\
&\le c|y| + (1+c)|y^*| = |y| - (1-c)|y| + (1+c)|y^*|.
\end{aligned}$$

and hypothesis (ii) is assured with $\beta = \frac{1+c}{1-c}|y^*|$.

Finally, we state a result on global asymptotic or exponential stability. In spirit, the following two results are closely related to Lu et al. (2007, Theorem 1). Note that we obtain a global result and are only concerned with fixed points, not general attractors. Also no assumption on the existence or invertibility of the Jacobian is required. In this sense the result is more general than those in Lu et al. (2007). Also strong ergodicity does not imply uniform asymptotic stability of the z-subsystem in the cascade (20), therefore the results in Lee and Jiang (2006), Loría and Nešić (2002), Loría and Nešić (2003) and Nešić and Loría (2004) are not applicable to the systems considered in the following theorem.

Theorem 3. *Let $\{P(k)\}_{k\in\mathbb{N}} \subset \mathbb{R}^{n\times n}$ be a strongly ergodic sequence of row stochastic matrices and assume G satisfies all conditions of Lemma 3. If y^* is a globally asymptotically stable fixed point of (21), then y^*e is a globally asymptotically stable fixed point for system (19).*

Proof. The assumptions of Theorem 1 are met and so it only remains to show global attractivity. Note that, by Lemma 3 all solutions of (19) are bounded. By Lemma 2 the ω-limit sets corresponding to all initial conditions lie in E. So consider an ω-limit set $\omega(x_0)$ and assume that $ye \in \omega(x_0)$ but $y \neq y^*$. Let U be a neighbourhood of y^*e on which local stability holds according to Theorem 1. We may assume $\mathrm{dist}(ye, U) > 0$. As $ye \in E$ it follows from Lemma 1 that all solutions $x(\cdot; k_0, ye)$ with the initial condition $x(k_0) = ye$ satisfy $\lim_{k\to\infty} x(k; k_0, ye) = y^*e$. Note that on E the system is time-invariant, so that there exists a time K, such that for all k_0 we have $x(k_0 + K; k_0, ye) \in U$. By assumption (i) the maps $x \mapsto P(k)x + G(x)e$ are equicontinuous on $B_\varepsilon(E)$ (i.e., each map, with respect to k, is uniformly continuous). Choose $\eta > 0$ such that

$$B_{\eta,\infty}(E) := \{x \in \mathbb{R}^n \; ; \; \mathrm{dist}_\infty(x, E) = \min_{r\in\mathbb{R}} \|x - re\|_\infty < \eta\}$$

is contained in $B_\varepsilon(E)$. The set $B_{\eta,\infty}(E)$ is forward invariant under all $F(k, \cdot)$, because if $\mathrm{dist}_\infty(x, E) = \|x - r_x e\|_\infty < \eta$, then as $\|P\|_\infty = 1$ for all row stochastic matrices

$$\mathrm{dist}_\infty(P(k)x + G(x)e, E) \leq \|P(k)(x - r_x e)\|_\infty < \eta \, .$$

Thus there exists a sufficiently small neighbourhood U_2 of ye such that for all $k_0 \in \mathbb{N}$ the solution corresponding to the initial condition $x(k_0) \in U_2$ satisfies $x(k_0 + K; k_0, x(k_0)) \in U$. But then by local stability, it follows that $x(k; k_0, x(k_0)) \in U$ for all $k \geq k_0 + K$. We thus arrive at a contradiction, if $ye \in \omega(x_0)$, then there exists a sequence $k_\ell \to \infty$ so that $\lim x(k_\ell; 0, x_0) = ye$. But then $x(k_\ell; 0, x_0) \in U_2$ for a sufficiently large ℓ and hence $x(k; 0, x_0) \in U$ for all $k \geq k_\ell + K$. Hence no subsequence of $\{x(k)\}$ converges to ye. This contradiction completes the proof. □

The previous result can be sharpened, if we assume exponential stability of the fixed point of (21). We omit the proof, which uses the same arguments as the proof of Theorem 3 and which could also be obtained easily by appealing to the methods used in Loría and Nešić (2002).

Theorem 4. *Let $\{P(k)\}_{k\in\mathbb{N}} \subset \mathbb{R}^{n\times n}$ be a uniformly exponentially ergodic sequence of row stochastic matrices and assume $G : \mathbb{R}^n \to \mathbb{R}$ is differentiable and satisfies conditions* (i) *and* (ii) *of Lemma 3. Then y^*e is globally uniformly exponentially stable for system (19).*

3.3. Switched systems

Given a compact set of row stochastic matrices $\mathcal{P} \subset \mathbb{R}^{n\times n}$, we may consider the switched system

$$x(k+1) = P(k)x(k) + G(x(k))e \, , \tag{29}$$

where $P(k) \in \mathcal{P}$. The results obtained so far have some immediate consequences for consensus under feedback with arbitrary switching. It is well-known that all sequences $\{P(k)\}_{k\in\mathbb{N}} \in \mathcal{P}^{\mathbb{N}}$ are strongly ergodic if and only if all sequences in $\mathcal{P}^{\mathbb{N}}$ are uniformly exponentially ergodic (Lu et al., 2007). In this case we call $\mathcal{P}$ uniformly ergodic. The rate of convergence towards E is in fact given by the projected joint spectral radius (Lu et al., 2007).

With this in mind the results obtained so far have immediate consequences for switched systems of the form (29). We note one of these consequences.

Corollary 1. *Let $\mathcal{P}$ be a compact set of row stochastic matrices that is uniformly ergodic and suppose that $G : \mathbb{R}^n \to \mathbb{R}$ is differentiable and satisfies conditions* (i) *and* (ii) *of Lemma 3. Then y^*e is globally uniformly exponentially stable for the switched system (29) under arbitrary switching.*

4. Optimised consensus for a speed advisory system

In this section, we describe an application to design a speed advisory system (SAS) for a fleet of vehicles. The objective of this system is to reduce CO_2 emissions of vehicles running on the highway. Full details of this application are given in Liu, Ordóñez-Hurtado, Wirth, Gu, Crisostomi, and Shorten (2016). Roughly speaking, we use the idea of optimised consensus, as described above, to calculate a set of virtual speeds, which the driver of each vehicle can use to guide an optimal travelling velocity[1] .

Let us consider a scenario in which a number of vehicles are driving along a given stretch of highway on different lanes in the same direction. Let N denote the total number of vehicles on a particular section of the highway where the SAS broadcast signal can be received. Each vehicle is equipped with a specific communication device, which is able to receive and transmit messages to other vehicles or the road infrastructure nearby, is regarded as a mobility agent. We define $\underline{N} := \{1, 2, \ldots, N\}$ for indexing the agents. Let $x_i(k)$ denote the recommended speed of the ith agent at time slot k. The corresponding recommended speed vector for all vehicles at time k is given by $x(k) := [x_1(k), x_2(k), \ldots, x_N(k)]^{\mathrm{T}}$. In addition, each agent is associated with a CO_2 emission cost function $f_i : \mathbb{R} \to \mathbb{R}$, which we assume to be convex, continuous and second order differentiable. We also assume that each agent can adjust $x_i(k)$ based on the knowledge of $f_i(x_i(k))$. The first derivative of f_i is denoted as $f_i' : \mathbb{R} \mapsto \mathbb{R}$. In our study, we shall adopt the average-speed model proposed in Boulter, Barlow, and McCrae (2009) to model each CO_2 emission cost function f_i over an interval of positive speeds, as a function of the average speed x_i by

$$f_i(x_i) = \mathrm{k}\left(\frac{\mathrm{a} + \mathrm{b}x_i + \mathrm{c}x_i^2 + \mathrm{d}x_i^3 + \mathrm{e}x_i^4 + \mathrm{f}x_i^5 + \mathrm{g}x_i^6}{x_i}\right), \tag{30}$$

where a, b, c, d, e, f, g, k $\in \mathbb{R}$ are used to specify different levels of emissions by different classes of vehicles — see Liu et al. (2016) for details. The specific mathematical problem is to find an optimal consensus point satisfying $x^* = y^*e$ such that the following optimisation problem is solved:

$$\min_{x\in\mathbb{R}^N} \quad \sum_{i=1}^{N} f_i(x_i) \quad \text{s.t. } x_i = x_j, \quad i, j \in \underline{N} \, . \tag{31}$$

After finding this common suggested speed, drivers are encouraged to drive at this recommended speed to minimise group emissions. Note in this study we use COPERT emission functions (Boulter et al., 2009), but it is also possible to measure these

[1] Note that each driver is still controlling their vehicle and the speeds are not adjusting automatically.

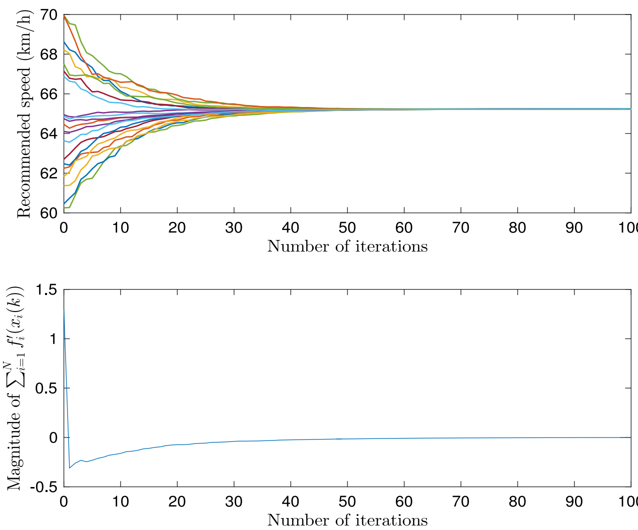

Fig. 1. Dynamics of the state variables $x(k)$ with $\mu = 0.1$.

average-speed functions in the car so as to incorporate individual driver behaviour. In what follows, we wish to use an iterative feedback scheme of the form (19) to solve the optimisation problem (31). We will require that this problem has a unique solution and derive the specific form for G in (19) from first order optimality conditions. To this end, it follows from elementary optimisation theory that when the f_i's are strictly convex, the optimisation problem will be solved if and only if there exists a unique $y^* \in \mathbb{R}$ such that $\sum_{i=1}^{N} f_i'(y^*) = 0$. With this in mind we apply a feedback signal $G(x) = -\mu \sum_{i=1}^{N} f_i'(x_i)$ where $\mu \in \mathbb{R}$ is a parameter to be determined. This gives rise to the following dynamical system

$$x(k+1) = P(k)x(k) - \mu \sum_{i=1}^{N} f_i'(x_i(k))e \tag{32}$$

where for each k we define $P(k)$ as

$$P_{ij}(k) = \begin{cases} 1 - \sum_{j \in N_k^i} \eta_j & \text{if } j = i \\ \eta_j & \text{if } j \in N_k^i \\ 0 & \text{otherwise} \end{cases} \tag{33}$$

where $\eta_j \in \mathbb{R}$ is a weighting factor, and N_k^i represents the set of neighbours communicating to the ith vehicle.

As we assume that the f_i are strictly convex, their derivatives are strictly increasing. We assume that, over the interval of interest, and for appropriate parameter values, each f_i' has a strictly positive and bounded growth, i.e., there exist constants $d_{\min}^{(i)}$ and $d_{\max}^{(i)}$; such that for any $a \neq b$

$$0 < d_{\min}^{(i)} \leq \frac{f_i'(a) - f_i'(b)}{a - b} \leq d_{\max}^{(i)} \qquad \forall i \in \underline{N}. \tag{34}$$

We claim that provided μ is chosen according to $0 < \mu < 2\left(\sum_{i=1}^{N} d_{\max}^{(i)}\right)^{-1}$ then (32) is uniformly globally asymptotically stable at the unique optimal point x^*e of the optimisation problem (31). First, we consider the scalar system associated with (32) which is given by

$$y(k+1) = y(k) - \mu \sum_{i=1}^{N} f_i'(y(k)). \tag{35}$$

Note first that the fixed point condition for (35) is $\sum_{i=1}^{N} f_i'(y^*) = 0$. So that a fixed point y^* of (35), gives rise, by Lemma 1 to a fixed point of (32), which satisfies the necessary and sufficient conditions for optimality. Now, we wish to use Theorem 3 to show global asymptotic stability. To this end, we need to verify that system (32) satisfies all the conditions of Theorem 3. The condition on μ ensures that the right hand side of (35) is in fact a strict contraction on $\mathbb{R}$. It follows from our comments after Lemma 3 that the assumption (ii) of Lemma 3 is satisfied. To show the Lipschitz condition (i) note that by (34) each f_i' is globally Lipschitz. As the coordinate functions are globally Lipschitz and sums of globally Lipschitz functions retain that property we obtain condition (ii).

To illustrate this application we consider 20 vehicles travelling along a section of road with random initial suggested speeds uniformly distributed between 60km/h and 70km/h. We assumed that there were 10 vehicles for each emission class, and the parameters set of the cost function for each class was chosen as R007 and R104, respectively, from Liu et al. (2016). The simulation results are presented in Fig. 1. Our results show that the recommended speeds of vehicles will asymptotically converge to the optimal one in less than 100 algorithm iterations.

5. Conclusion

In this note we present a rigorous proof of stability and convergence of a recently proposed consensus system with feedback. Examples are given to illustrate the usefulness of the algorithm.

References

Blondel, V. D., Hendrickx, J. M., Olshevsky, A., & Tsitsiklis, J. N. (2005). Convergence in multiagent coordination, consensus, and flocking. In *Proc. 44th IEEE conf. decision and control and Eur. control conf. (CDC-ECC '05)* (pp. 2996–3000).

Boulter, P. G., Barlow, T. J., & McCrae, I. S. (2009). *Emission factors 2009: Report 3-exhaust emission factors for road vehicles in the United Kingdom. TRL Report PPR356*. Wokingham: TRL Limited.

Burger, M., Notarstefano, G., & Allgower, F. (2014). A polyhedral approximation framework for convex and robust distributed optimization. *IEEE Trans. Automat. Control*, *59*(2), 384–395.

Chatterjee, S., & Seneta, E. (1977). Towards consensus: some convergence theorems on repeated averaging. *J. Appl. Probab.*, *14*(1), 89–97.

Ipsen, I. C. F., & Selee, T. M. (2011). Ergodicity coefficients defined by vector norms. *SIAM J. Matrix Anal. Appl.*, *32*(1), 153–200.

Knorn, F., Corless, M., & Shorten, R. (2011). Results in cooperative control and implicit consensus. *Internat. J. Control*, *84*(3), 476–495.

Knorn, F., Stanojevic, R., Corless, M., & Shorten, R. (2009). A framework for decentralised feedback connectivity control with application to sensor network. *Internat. J. Control*, *82*(11), 2095–2114.

Lee, T. C., & Jiang, Z. P. (2006). On uniform global asymptotic stability of nonlinear discrete-time systems with applications. *IEEE Trans. Automat. Control*, *51*(10), 1644–1660.

Liu, M., Ordóñez-Hurtado, R. H., Wirth, F., Gu, Y., Crisostomi, E., & Shorten, R. (2016). A distributed and privacy-aware speed advisory system for optimizing conventional and electric vehicle networks. *IEEE Transactions on Intelligent Transportation Systems*, *17*(5), 1308–1318.

Loría, A., & Nešić, D. (2002). Stability of time-varying discrete-time cascades. In *Proc. 15th. IFAC world congress*, 35 (1), 263–268.

Loría, A., & Nešić, D. (2003). On uniform boundedness of parameterized discrete-time systems with decaying inputs: Applications to cascades. *Systems Control Lett.*, *49*(3), 163–174.

Lu, W., Atay, F. M., & Jost, J. (2007). Synchronization of discrete-time dynamical networks with time-varying couplings. *SIAM J. Math. Anal.*, *39*(4), 1231–1259.

Mesbahi, M., & Egerstedt, M. (2010). *Graph theoretic methods in multiagent networks*. Princeton University Press.

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

Nedic, A., Ozdaglar, A., & Parrilo, P. A. (2010). Constrained consensus and optimization in multi-agent networks. *IEEE Trans. Automat. Control*, *55*(4), 922–938.

Nešić, D., & Loría, A. (2004). On uniform asymptotic stability of time-varying parameterized discrete-time cascades. *IEEE Trans. Automat. Control*, *49*(6), 875–887.

Olfati-Saber, R., Fax, J. A., & Murray, R. M. (2007). Consensus and cooperation in networked multi-agent systems. *Proc. IEEE*, *95*(1), 215–233.

Shi, G., Johansson, K. H., & Hong, Y. (2013). Reaching an optimal consensus: dynamical systems that compute intersections of convex sets. *IEEE Trans. Automat. Control*, *58*(3), 610–622.

Mingming Liu received his B.E. and Ph.D. degrees from National University of Ireland Maynooth in 2011 and 2015 respectively. His research interests are nonlinear system dynamics, distributed control techniques, modelling and optimisation in the context of smart grid, and smart transportation systems.

Fabian Wirth is a Professor at the University of Passau, Germany. His current interests include stability theory, queueing networks, switched systems and large scale networks with applications to networked systems and in the domain of smart cities.

Martin Corless is a Professor in the School of Aeronautics & Astronautics at Purdue University. His research is concerned with obtaining tools which are useful in the robust analysis and control of systems containing significant uncertainty and in applying these results to aerospace and mechanical systems and to sensor and communication networks.

Robert Shorten is currently Professor of Control Engineering and Decision Science at University College Dublin. He has been active in computer networking, automotive research, collaborative mobility (including smart transportation and electric vehicles), as well as basic control theory and linear algebra.