Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 49-18 (2016) 552–557

A Distributed Algorithm for Computing a Common Fixed Point of a Family of Paracontractions

Daniel Fullmer * **Lili Wang** * **A. Stephen Morse** *

* *Department of Electrical Engineering, New Haven, CT 06520 USA (e-mail: {daniel.fullmer, lili.wang, as.morse}@yale.edu)*

Abstract:
A distributed algorithm is described for finding a common fixed point of a family of $m > 1$ nonlinear maps $M_i : \mathbb{R}^n \to \mathbb{R}^n$ assuming that each map is a paracontraction and that such a common fixed point exists. The common fixed point is simultaneously computed by m agents assuming each agent i knows only M_i, the current estimates of the fixed point generated by its neighbors, and nothing more. Each agent recursively updates its estimate of the fixed point by utilizing the current estimates generated by each of its neighbors. Neighbor relations are characterized by a time-dependent directed graph $\mathbb{N}(t)$ whose vertices correspond to agents and whose arcs depict neighbor relations. It is shown that for any family of paracontractions $M_i,\ i \in \{1, 2, \ldots, m\}$ which has at least one common fixed point, and any sequence of strongly connected neighbor graphs $\mathbb{N}(t)$, $t = 1, 2, \ldots$, the algorithm causes all agent estimates to converge to a common fixed point.

Keywords: distributed algorithm, paracontraction, nonlinear

1. INTRODUCTION

This paper is concerned with distributed algorithms for enabling a group of $m > 1$ mobile autonomous agents to solve certain types of nonlinear equations over a network. It is assumed that each agent can receive information from its neighbors where by a *neighbor* of agent i is meant any other agent within agent i's reception range. We write $\mathcal{N}_i(t)$ for the labels of agent i's neighbors at time t, and we always take agent i to be a neighbor of itself. Neighbor relations at time t can be conveniently characterized by a directed graph $\mathbb{N}(t)$ with m vertices and a set of arcs defined so that there is an arc in $\mathbb{N}(t)$ from vertex j to vertex i just in case agent j is a neighbor of agent i at time t. Each agent i has a real-time dependent state vector $x_i(t)$ taking values in $\mathbb{R}^n$, and we assume that the information agent i receives from neighbor j at time t is $x_j(t)$. It is also assumed that agent i knows a suitably defined nonlinear map $M_i : \mathbb{R}^n \to \mathbb{R}^n$ and that all of the M_i share at least one common fixed point. In general terms, the problem of interest is to develop algorithms, one for each agent, which will enable all m agents to iteratively compute a common fixed point of all of the M_i.

Motivation for this problem stems, in part, from Mou et al. (2015) which deals with the problem of devising a distributed algorithm for finding a solution to the linear equation $Ax = b$, assuming the equation has at least one solution, and agent i knows a pair of the matrices $(A_i^{n_i \times n}, b_i^{n_i \times 1})$ where $A = \begin{bmatrix} A_1' & A_2' & \cdots & A_m' \end{bmatrix}'$ and $b = \begin{bmatrix} b_1' & b_2' & \cdots & b_m' \end{bmatrix}'$. Assuming each A_i has linearly independent rows, one local update rule for solving this problem is of the form

$$x_i(t+1) = L_i(z_i(t))$$

where $L_i : \mathbb{R}^n \to \mathbb{R}^n$ is the affine linear map $x \longmapsto x - A_i'(A_i A_i')^{-1}(A_i x - b_i)$,

$$z_i(t) = \frac{1}{m_i(t)} \sum_{j \in \mathcal{N}_i(t)} x_j(t),$$

and $m_i(t)$ is the number of labels in $\mathcal{N}_i(t)$ (Wang et al. (2016)). The map L_i is an example of a 'paracontraction' with respect to the two norm on $\mathbb{R}^n$. More generally, a continuous nonlinear map $M : \mathbb{R}^n \to \mathbb{R}^n$ is a *paracontraction* with respect to a given norm $\|\cdot\|$ on $\mathbb{R}^n$, if $\|M(x) - y\| < \|x - y\|$ for all $x \in \mathbb{R}^n$ satisfying $x \neq M(x)$ and all $y \in \mathbb{R}^n$ satisfying $y = M(y)$ (Elsner et al. (1992)). One obvious consequence of this definition is that $\|M(x) - y\| \leq \|x - y\|$ for all $x \in \mathbb{R}^n$ and all $y \in \mathbb{R}^n$ satisfying $y = M(y)$. Note that $y = L_i(y)$ if and only if $A_i y = b_i$ and for any such y, $L_i(x) - y = P_i(x - y)$ where P_i is the orthogonal projection matrix $P_i = I - A_i'(A_i A_i')^{-1} A_i$. Since the induced 2-norm of P_i is 1, $\|P_i(x - y)\|_2 \leq \|x - y\|_2,\ \forall x, y \in \mathbb{R}^n$ so $\|L_i(x) - y\|_2 \leq \|x - y\|_2,\ \forall x, y \in \mathbb{R}^n$. Moreover for any y satisfying $L_i(y) = y$, the inequality $x \neq L_i(x)$ is equivalent to $x - y \notin \ker A_i$ and $\ker A_i = \text{image } P_i$ so $x - y \notin \text{image } P_i$ whenever $x \neq L_i(x)$ and $y \in \text{image } P_i$. But for such x and y, $\|P_i(x - y)\|_2 < \|x - y\|_2$ so $\|L(x) - y\|_2 < \|x - y\|_2$. Clearly L_i is a paracontraction as claimed.

There are many other examples of paracontractions discussed in the literature. Some can be found in Elsner et al. (1992) and Byrne (2007). Here are several others.

* The problem to which this paper is addressed was prompted by useful discussions with Bronislaw Jakubczyk. This work was supported by the US Air Force Office of Scientific Research and by the National Science Foundation.

2405-8963
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.10.223

(1) The orthogonal projector $x \mapsto \arg\min_{y\in\mathcal{C}} \|x-y\|_2$ associated with a nonempty closed convex set $\mathcal{C}$. This been used for a number of applications including the constrained consensus problem in Nedić et al. (2010). The fixed points of this map are vectors in $\mathcal{C}$. (Elsner et al. (1992))
(2) The gradient map $x \longmapsto x - \alpha\nabla f(x)$ where $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is convex and differentiable, ∇f is Lipschitz continuous with parameter $\lambda > 0$, and α is a constant satisfying $0 < \alpha < \frac{2}{\lambda}$. The fixed points of this map are vectors in $\mathbb{R}^n$ which minimize f.
(3) The proximal map associated with a closed proper convex function $f : \mathbb{R}^n \to (-\infty, \infty]$. The fixed points of this map are vectors in $\mathbb{R}^n$ which minimize f. See Eckstein and Bertsekas (1992) as well as Parikh and Boyd (2014).

Paracontractions are also discussed in Xiao et al. (2006) and Wu (2007). What is especially important about paracontractions, whether they are linear or not, is the following well-known theorem published in Elsner et al. (1992).

Theorem 1. Let $M_1, M_2, \ldots, M_m$, be a finite set of m paracontractions with respect to any given norm on $\mathbb{R}^n$. Suppose that all of the paracontractions share at least one common fixed point. Let $\sigma(t),\ t \in \{1, 2, \ldots\}$ be an infinite sequence of integers from the set $\{1, 2, \ldots, m\}$ with the property that each integer in $\{1, 2, \ldots, m\}$ occurs in the sequence infinitely often. Then the state $x(t)$ of the iteration

$$x(t+1) = M_{\sigma(t)}(x(t)), \quad t \in \{1, 2, \ldots\}$$

converges to a common fixed point of the m paracontractions.

In the sequel we will use this result to establish the convergence of a family of distributed paracontracting iterations.

2. THE PROBLEM

The specific problem to which this paper is addressed is this. Let $M_1, M_2, \ldots, M_m$ be a set of m paracontractions with respect to the standard p-norm $\|\cdot\|$ on $\mathbb{R}^n$ where p is a constant satisfying $1 < p < \infty$. Suppose that all of the paracontractions share at least one common fixed point. Find conditions on the time-varying neighbor graph $\mathbb{N}(t)$ so that the states of all m iterations

$$x_i(t+1) = M_i\left(\frac{1}{m_i(t)} \sum_{j\in\mathcal{N}_i(t)} x_j(t)\right), \quad i \in \mathbf{m},\ t \geq 0 \quad (1)$$

converge to a common fixed point of the M_i where $\mathbf{m} \triangleq \{1, \ldots, m\}$ and $\mathcal{N}_i(t)$ is the set of labels of those agents which are neighbors of agent i at time t. The main result of this paper is as follows.

Theorem 2. If each of the neighbor graphs in the sequence $\mathbb{N}(1), \mathbb{N}(2), \ldots$ is strongly connected and the paracontractions $M_1, M_2, \ldots, M_m$ share at least one common fixed point, then the states $x_i(t)$ of the m iterations defined by (1), all converge to a common fixed point of the M_i as $t \to \infty$.

The remainder of this paper is devoted to a proof of this theorem.

3. ANALYSIS

To proceed, let us note that the family of m iterations given by (1) can be written as a single iteration of the form

$$x(t+1) = M((F(t) \otimes I)x(t)), \quad t \geq 0 \quad (2)$$

where for any set of vectors $\{x_i \in \mathbb{R}^n,\ i \in \mathbf{m}\}$, $x \in \mathbb{R}^{mn}$ is the stacked vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \quad (3)$$

$M : \mathbb{R}^{mn} \to \mathbb{R}^{mn}$ is the map

$$x \longmapsto \begin{bmatrix} M_1(x_1) \\ M_2(x_2) \\ \vdots \\ M_m(x_m) \end{bmatrix},$$

$F(t)$ is the $m \times m$ flocking matrix [1] determined by $\mathbb{N}(t)$, I is the $n \times n$ identity matrix, and $F(t) \otimes I$ is the Kronecker product of $F(t)$ with I.

It will be convenient to introduce the "average" vectors

$$z_i(t) = \frac{1}{m_i(t)} \sum_{j\in\mathcal{N}_i(t)} x_j(t), \quad i \in \mathbf{m},\ t \geq 0 \quad (4)$$

in which case the stacked vector

$$z(t) = \begin{bmatrix} z_1'(t) & z_2'(t) & \cdots & z_m'(t) \end{bmatrix}'$$

satisfies

$$z(t) = (F(t) \otimes I)x(t), \quad t \geq 0 \quad (5)$$

and consequently

$$z(t+1) = (F(t+1) \otimes I)M(z(t)), \quad t \geq 0 \quad (6)$$

because of (2). It is clear that convergence of all of the x_i to a single point in $\mathbb{R}^n$ implies convergence of all of the z_i to the same point. On the other hand, if all of the z_i converge to a single point which is, in addition, a common fixed point of the $M_i,\ i \in \mathbf{m}$, then because the M_i are continuous and $x_i(t+1) = M_i(z_i(t)),\ t \geq 0$, all of the x_i must converge to the same fixed point. In other words, convergence of all of the z_i to a common fixed point of the $M_i,\ i \in \mathbf{m}$, is equivalent to convergence of all of the x_i to the same fixed point. Thus to prove Theorem 2 it is enough to show that if all of the $\mathbb{N}(t)$ are strongly connected, the $z_i(t)$ all converge to a common fixed point y^* of the $M_i,\ i \in \mathbf{m}$.

It is obvious from (6) that for any positive integer q,

$$z(q) = ((F(q) \otimes I)M \circ \cdots \circ (F(1) \otimes I)M)(z(0)) \quad (7)$$

Prompted by this we will study the properties of maps from $\mathbb{R}^{mn}$ to $\mathbb{R}^{mn}$ which are of the form $x \longmapsto ((S(q) \otimes I)M \circ \cdots \circ (S(1) \otimes I)M)(x)$ where q is a positive integer, and $S(t),\ t \in \mathbf{q} \triangleq \{1, 2, \ldots, q\}$ is a family of q stochastic matrices $S(t) = [s_{ij}(t)]_{m\times m}$. We will show that under suitable conditions, such maps are paracontractions with respect to the *mixed vector norm* $\|\cdot\|_{p,\infty}$ on $\mathbb{R}^{mn}$ where

[1] By the *flocking matrix* of a neighbor graph $\mathbb{N}$ is meant that stochastic matrix $F = D^{-1}A'$ where A is the adjacency matrix of $\mathbb{N}$, D is a diagonal matrix whose ith diagonal entry is the in-degree of vertex i in $\mathbb{N}$, and prime denotes transpose.