arXiv:0803.2211v2 [math.OC] 9 May 2010

# On Conditions for Convergence to Consensus

Jan Lorenz, Dirk A. Lorenz

***Abstract*—A new theorem on conditions for convergence to consensus of a multiagent time-dependent time-discrete dynamical system is presented. The theorem is build up on the notion of averaging maps. We compare this theorem to results by Moreau (IEEE Transactions on Automatic Control, vol. 50, no. 2, 2005) about set-valued Lyapunov theory and convergence under switching communication topologies. We give examples that point out differences of approaches including examples where Moreau's theorem is not applicable but ours is. Further on, we give examples that demonstrate that the theory of convergence to consensus is still not complete.**

***Index Terms*—consensus protocol, averaging map, set-valued Lyapunov theory, multiagent systems.**

## I. Introduction

In this technical note we analyze discrete dynamical systems of consensus formation as presented in the context of distributed computing [1], [2], flocking (e.g. of unmanned aerial vehicles) [3]–[5] and general as multi-agent coordination problems [6]–[8] (to mention just a few). The dynamical system may also be called 'agreement algorithm' or 'consensus protocol'. The convergence theorems of Moreau [6] together with the extensions of Angeli and Bliman [9] are the most general ones. The main theorem of Moreau states conditions for convergence to consensus under switching communication topologies. Convergence to consensus is there implied by 'global asymptotic stability of the set of equilibrium solutions with consensus as equilibrium points'. Conditions are on the one hand on the communication topologies in their time-evolution and on the other hand on the updating maps. Moreau applied a set-valued Lyapunov theory, which uses a set-valued function on the state space which is contractive with respect to the updating map. This implies convergence of the set to a singleton.

We contribute a similar but new approach based on the notion of an averaging map. Moreau deals with communication topologies by defining conditions on how many successive communication topologies must be regarded until the composition of these updating maps fulfills the contraction properties used to apply the set-valued Lyapunov theory. We skip the issue on changing communication topologies and deal directly with maps which fulfill a contraction property which is different from Moreau's.

Our theorem generalizes a result of Krause [10] by allowing arbitrary switching between different averaging maps but follow the same line of compactness, continuity and convexity arguments.

Section II presents the convergence result and possible extensions. Section III discusses the relations to two of Moreau's theorems in more detail. Section IV gives examples and counterexamples to show existing gaps in the theory of consensus algorithms. All proofs of lemmas and theorems are collected in Appendix A.

J. Lorenz is with the Chair of Systems Design, ETH Zurich, Kreuzplatz 5, 8032 Zurich, Switzerland, most of the work was done when he was with the Department of Mathematics and Computer Science, University of Bremen, Bibliothekstrasse 1, 28359 Bremen, Germany, post@janlo.de, http://www.janlo.de.

D.A. Lorenz is with the Institute for Analysis and Algebra, Carl-Friedrich Gauß Department, TU Braunschweig, 38092 Braunschweig, Germany, d.lorenz@tu-braunschweig.de, http://www.tu-braunschweig.de/iaa/personal/lorenz.

Manuscript received xxx 00, 2008; revised xxx 00, 2008.

## II. Convergence result

We consider a dynamical system of the form

$$x(t+1) = f_t(x(t)) \tag{1}$$

with discrete time $t \in \mathbb{N}$. Dynamics take place in a $d \times n$-dimensional space: We consider a set of agents $\underline{n} = \{1, \dots, n\}$ where each of them has coordinates in a $d$-dimensional set $S \subset \mathbb{R}^d$. Hence, the solutions of (1) have the form $x : \mathbb{N} \to S^n \subset \mathbb{R}^{d\times n}$. The individual coordinates of agent $i$ at time $t \in \mathbb{N}$ is labeled $x^i(t) \in S$, and $x(t) \in S^n$ is called the *profile* at time $t \in \mathbb{N}$. Finally, the mappings $f_t$ which govern the dynamics are of the form $f_t : S^n \to S^n$. We denote the component functions by $f_t^i$.

To state our main result on convergence of such systems to consensus we introduce the following notations. An element $x \in S^n$ is called *consensus* if all $d$-dimensional coordinates $x^i$ have the same value, i.e. there exists a vector $\gamma \in S$ such that $x^i = \gamma$ for $i \in \underline{n}$. By $\operatorname{conv}_{i\in\underline{n}} x^i$ we define the convex hull of the vectors $x^1, \dots, x^n$.

The core notion in this note is an 'averaging map'. We build the definition of an averaging map on a generalized convex hull. Consider a continuous function $y : S^n \to S^m$ which maps a profile to a certain set of $m$ vectors $y(x) = (y^1(x), \dots, y^m(x))$ such that for all $x \in S^n$ and all $i \in \underline{n}$ it holds $x^i \in \operatorname{conv}_{j\in\underline{m}} y^j(x)$. We call such a function $y$ a generalized *barycentric coordinate map* and we call $\operatorname{conv}_{j\in\underline{m}} y^j(x)$ the *$y$-convex hull* of the vectors $x^1, \dots, x^n$. (We call $y$ 'generalized' because it needs not be a bijective transformation.) So, a $y$-convex hull is a set-valued function from $S^n$ to the compact and convex subsets of $S$. We call a set $y$-convex, if it is the union of the $y$-convex hulls of all $n$ of its points. Examples for $y$-convex hulls include the convex hull itself, and the multidimensional interval $[\min_{i\in\underline{n}} x^i, \max_{i\in\underline{n}} x^i]$ (with $\min$ and $\max$ applied componentwise). For the first it holds $m = n$ for the second $m = 2^d$. Many other examples fit into this setting: the smallest interval for any basis of $\mathbb{R}^d$ [9, Example 2], or smallest polytope with faces parallel to a set of $k \geq d+1$ hyperplanes [9, Example 3] containing $x^1, \dots, x^n$ (the generalized barycentric coordinates are then the extreme points of the polytope, perhaps with multiples to have a constant $m$). Now, we define the central notion of this paper.

*Definition 2.1:* Let $S \subset \mathbb{R}^d$, $y : S^n \to S^m$ be a generalized barycentric coordinate map such that $S$ is $y$-convex. A mapping $f : S^n \to S^n$ is called a *$y$-averaging map*, if for every $x \in S^n$ it holds

$$\operatorname*{conv}_{i\in\underline{m}} y^i(f(x)) \subset \operatorname*{conv}_{i\in\underline{m}} y^i(x). \tag{2}$$

Furthermore, a *proper $y$-averaging map* is a $y$-averaging map, such that for every $x \in S^n$ which is not a consensus, the above inclusion is strict.

A $y$-averaging map maps a profile $x$ into its $y$-convex hull. Furthermore, the $y$-convex hull of the new profile $f(x)$ lies in the $y$-convex hull of the vectors $x^1, \dots, x^n$. Hence, we may also work with the $y$-convex hull of the initial profile $x(0)$ instead of the set $S$. Sometimes it is useful to look at the contraposition of the definition of proper: If equality holds in (2) this implies that $x$ is a consensus. In the following we may omit '$y$' when we mention an averaging map, but for an averaging map the definition of $y$ is a prerequisite. The best proxy for the mind is $y = \mathrm{id}$.

Since we are going to consider families of averaging maps we introduce the concept of equiproper averaging maps. To this end, we need the Hausdorff distance on the set of compact subsets of a metric space $(X, d)$. The distance of a point $x \in X$ and a nonempty compact set $C \subset X$ is defined as $d(x, C) := \min_{c\in C} d(x, c)$. Let $B, C \subset X$ be nonempty and compact, then the *Hausdorff distance* is defined as

$$d_H(B,C) := \max\{\max_{b\in B} d(b,C), \max_{c\in C} d(c,B)\}.$$

Equivalently, one can say that the Hausdorff distance is the smallest $\varepsilon$ such that the $\varepsilon$-neighborhood of $B$ contains $C$ and the $\varepsilon$-neighborhood of $C$ contains $B$. It is easy to see that $d_H(B,C) = 0$ holds if and only if $B = C$. In the special case $B \subset C \subset S \subset \mathbb{R}^d$ it holds

$$d_H(B,C) = \max_{b\in B} d(b,C) = \max_{b\in B}\min_{c\in C} \|b-c\|\,. \tag{3}$$

*Definition 2.2:* Let $y$ be a generalized barycentric coordinate map and let $F$ be a family of proper $y$-averaging maps. $F$ is called *equiproper*, if for every $x \in S^n$ which is not a consensus, there is $\delta(x) > 0$ such that for all $f \in F$

$$d_H\big(\operatorname*{conv}_{i\in\underline{m}} y^i(f(x)), \operatorname*{conv}_{i\in\underline{m}} y^i(x)\big) > \delta(x). \tag{4}$$

Now we state a lemma which says that the family of equiproper $y$-averaging maps is closed under pointwise limits.

*Lemma 2.3:* Let $f_t$ be a sequence of $y$-averaging maps forming an equiproper family of $y$-averaging maps such that $f_t \to g$ pointwise. Then $g$ is a proper $y$-averaging map.

Now we are able to state our main theorem.

*Theorem 2.4:* Let $S \subset \mathbb{R}^d$, $y$ be a generalized barycentric coordinate map such that $S$ is $y$-convex, and $F$ be an equicontinuous family of equiproper $y$-averaging maps on $S^n$. Then it holds for any sequence $(f_t)_{t\in\mathbb{N}}$ with $f_t \in F$ and any $x(0) \in S^n$ that the solution of (1) converges to a consensus, i.e. there exists $\gamma \in S$ such that for all $i \in \underline{n}$ it holds $\lim_{t\to\infty} x^i(t) = \gamma$.

Notice that the limit $\gamma$ depends not only on the initial value $x(0)$ but also on the realization of the sequence $(f_t)_{t\in\mathbb{N}}$, however, $\gamma$ depends continuously on the intial value if the sequence $(f_t)$ is fixed as the following lemma and corollary show.

*Lemma 2.5:* Let $(X,d)$ be a metric space and $f_t : X \to X$ be such that the solution of $x(t+1) = f_t(x(t))$ converge to some limit for every initial value $x(0) \in X$. Then the limit depends continuously on the initial value if $\{f_t\}$ is an equicontinuous family.

The following corollary is a direct consequence.

*Corollary 2.6:* Let the sequence $(f_t)$ in the situation of Theorem 2.4 be fixed. Then the consensus value $\gamma$ (which exists due to Theorem 2.4) depends continuously on the initial value.

Theorem 2.4 is a generalization of a theorem of Krause [10]. Krause's theorem is the special case when $y$ is the identity and $F$ contains only one proper averaging map. Notice that 'equi' in equiproper and equicontinuous can be omitted if $F$ is a finite set. An easy extension is to allow $F$ to contain also non-proper averaging maps (but at least one proper averaging map). Then the sequence $(f_t)_{t\in\mathbb{N}}$ has to contain a subsequence $(f_{t_s})_{s\in\mathbb{N}}$ of equiproper averaging maps to ensure convergence to consensus. This holds because then $\{g_s \,|\, g_s = f_{t_s}\circ\cdots\circ f_{t_{s+1}}\}$ is an equiproper set of averaging maps for $s \in \mathbb{N}$. Notice that it is possible that a sequence of averaging maps contains a subsequence as above such that subcompositions $g_s$ form an equiproper set, even when no $f_t$ is proper. The easiest example is when $F$ contains only one linear map which is determined by a row-stochastic square matrix which is regular but not scrambling (see Seneta [11]). For linear systems 'row-stochastic' is equivalent to 'being an averaging map' (with $y$ the identity) and 'scrambling' is equivalent to 'proper'. From the theory of nonnegative matrices we know that for each regular matrix there is an integer such that higher powers are scrambling.

In the spirit of [9] we state another generalization of Theorem 2.4 which deals with deformations of the hull. To this end, let $S, T \subset \mathbb{R}^d$ be compact and $\phi : T \to S$ be a homeomorphism. For a generalized barycentric coordinate map $y : S^n \to S^m$ we define the $y,\phi$-hull as $\phi^{-1}(\mathrm{conv}_{i\in\underline{m}}\, y^i(\phi(x)))$. Now, a $y,\phi$-averaging map $g$ is defined analogous to Definition 2.1:

$$\phi^{-1}(\operatorname*{conv}_{i\in\underline{m}} y^i(\phi(g(x)))) \subset \phi^{-1}(\operatorname*{conv}_{i\in\underline{m}} y^i(\phi(x))).$$

Note, that the $y,\phi$-hull is not necessarily convex, see [9, Example 6]. The extension of the notions 'proper' and 'equiproper' is straightforward.

*Theorem 2.7:* Let $\phi : T \to S$ be continuous with Lipschitz continuous inverse and let $y$ be a generalized barycentric coordinate map such that $S$ is $y$-convex. Let $G$ be a family of equicontinuous, equiproper $y,\phi$-averaging maps on $T^n$. Then it holds for any sequence $(g_t)_{t\in\mathbb{N}}$ with $g_t \in G$ and any $x(0) \in T^n$ that the solution of $x(t+1) = g_t(x(t))$ converges to a consensus.

## III. Comparison with Moreau's set-valued Lyapunov theory and main theorem

Theorem 2.4 has similarities to Moreau's set-valued Lyapunov Theorem [6, Theorem 4]. This theorem implies global asymptotic stability of the set of equilibrium solutions when there exists a set-valued function $V$ on the state space, a measure for these sets $\mu$, and a positive definite function $\beta$ on the state space. Essentially it has to hold $V(f_t(x)) \subset V(x)$ and $\mu(V(f_t(x))) - \mu(V(x)) \leq -\beta(x)$. The best example to imagine is $V = \mathrm{conv}$, and $\mu$ is the diameter of a set.

The set of equilibrium solutions for the dynamical system (1) under the conditions of Theorem 2.4 contains only all constant solutions on consensus vectors, due to the equiproperness of $F$. Given this set of equilibrium solutions, "global asymptotic stability of the set of equilibrium solutions" implies convergence to consensus. Convergence to consensus is thus a special case of the set-valued Lyapunov Theorem in [6]. To the best of our knowledge, it is the only case in which the theorem has been used so far.

Compared with our Theorem 2.4 the role of the set-valued map $V$ is taken by the $y$-convex hull. So, we also deal with a general class of functions due to the various possible coordinate maps $y : S^n \to S^m$—we only assume that $m$ is finite. However, we do not need a general measure $\mu$ on these maps. The assumption $\mu(V(f_t(x))) - \mu(V(x)) < \beta(x)$ corresponds to $d_H(V(f_t(x)), V(x)) > \delta(x)$. This is a different condition and often weaker, as for example in the case where Moreau specifies it to proof his main Theorem [6, Theorem 2]. There $\mu$ is the diameter of $V(x)$ (which he specifies as the $\mathrm{conv}(x)$).

Theorem 2.4 has also similarities to Moreau's main theorem [6, Theorem 2]. This theorem is more specific than Theorem 2.4 by incorporating switching communication topologies. Its main drawback is that it relies very much on convex hulls (see [9] for a method to overcome this drawback). Our result generalizes to convex hulls of generalized coordinate maps. Further on, in Moreau's theorem agents are forced to move into the relative interior of the convex hull (respecting the communication topology). Specifically, this implies that agents have to leave all extreme points of the convex hull (of agents in its neighborhood) after one iteration. Our theorem needs only agents at one arbitrary extreme point (of the global $y$-convex hull) to leave it towards the interior after one iteration. This is implied by properness of averaging maps. The assumption 'equiproper' in our theorem finds its analog in Moreau's theorem by assuming that the sets $e_k(\mathcal{A}(t))(x)$ are chosen independently of $t$.[1]

Summarizing the above one can say that both Moreau's theorem and Theorem 2.4 are similar. However, the assumptions as well as the methods of proof are different. On the one hand we do not incorporate switching communication topologies explicitly, but on the other hand we need weaker conditions for the updating maps $f_t$. Further on, we generalized to $y$-convex hulls and are also able to incorporate the extensions of Moreau's theorem by Angeli and Bliman [9] to overcome the restriction to convex sets. Moreover, the notion of a (equi-)proper $y$-averaging map allows a systematic and structured treatment of consensus algorithms (see e.g. the results in Lemma 2.3 and Lemma 2.5). Hence, Theorem 2.4 together with 2.7 give an alternative approach to the analysis of consensus protocols whose applicability is illustrated by examples in the next section.

## IV. Examples and Counterexamples

In this section we present counterexamples (Examples 4.1–4.3) to point that the existing theory, including our Theorem 2.4, delivers no sharp results on convergence to consensus. We also give examples which show cases, where our theorem is applicable but Theorem 2 of Moreau is not (Examples 4.4–4.6).

Continuity, for instance, is not necessary for convergence to consensus since there are discontinuous proper averaging maps which converge to consensus (one may take different averaging maps on different subdomains of $S$). On the other hand discontinuity may destroy convergence to consensus even for proper averaging maps (see [12, Section 3.1] for examples for this phenomenon).

The next two examples illustrate the role of equiproperness.

*Example 4.1 (Non-equiproper not leading to consensus):* Let

$$f_t(x^1,x^2) := \left((1-\frac{1}{4^t})x^1 + \frac{1}{4^t}x^2\ ,\ \frac{1}{4^t}x^1 + (1-\frac{1}{4^t})x^2\right)$$

It is easy to see that for $t \geq 1$ and $x(1) = (0,1)$ it holds that $x^1(t) < \frac{1}{3}$ and $x^2(t) > \frac{2}{3}$. Obviously, $\{f_t \mid t \in \mathbb{N}\}$ is not equiproper because $f_t$ converges to the identity as $t \to \infty$.

*Example 4.2 (Non-equiproper leading to consensus):* Let

$$f_t(x^1,x^2) := \left((1-\frac{1}{t})x^1 + \frac{1}{t}x^2\ ,\ x^2\right)$$

This example is not equiproper, because $f_t$ converges to the identity for $t \to \infty$. Thus, Theorem 2.4 does not apply, but for $t \geq 2$ and any $x(2) \in (\mathbb{R})^2$ the system $x(t+1) = f_t(x(t))$ has the solution $x(t) = (\frac{1}{t-1}x^1(2) + \frac{t-2}{t-1}x^2(2)\ ,\ x^2(2))$ and thus converges to consensus at $x^2(2)$. Note that the convergence is not at an exponential rate.
Convergence to consensus in the last example can also not be ensured by Moreau's theorems.

The next example illustrates the role of equicontinuity and is inspired by bounded confidence [13].

*Example 4.3 (Vanishing confidence):* Let $f_t : \mathbb{R}^n \to \mathbb{R}^n$ with

$$f_t^i(x) := \frac{\sum_{j=1}^n D_t(|x^i - x^j|)x^j}{\sum_{j=1}^n D_t(|x^i - x^j|)}$$

and $D_t : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. Now, $f_t$ is an averaging map for any choice of $D_t$. Further on, $f_t$ is continuous if $D_t$ is, and $f_t$ is proper if $D_t$ is strictly positive. We chose $D_t(y) := e^{-(\frac{y}{\varepsilon})^t}$ as a sequence of functions which has the cutoff function as pointwise limit function. Hence, $D_t$ is continuous but $\{D_t \mid t \in \mathbb{N}\}$ is not equicontinuous. For $x(0) = (0,8)$, $\varepsilon = 1$ the process $x(t) = f_t(x(t))$ does not converge to consensus although only proper averaging maps are involved. Rough estimates show that $\left|x^1(t) - x^2(t)\right| \geq 4$.

For other settings convergence under vanishing confidence is possible, as numerical examples in [12] show.

The following examples are to show limitations of Moreau's Theorem 2 and how Theorem 2.4 can be applied to show convergence to consensus.

*Example 4.4 (Rendezvous problem with watergun sensors):* We consider a version of the Rendezvous Problem [14] where $n$ agents are to locate themselves decentralized at the same position in twodimensional space. Each agent has three waterguns, an *activation gun* and two *search guns*. Agents can perceive from which kind of gun they were hit and can respond (e.g. acoustically). The search gun is used as a sensor to check if there is at least one other agent in direction $\alpha \in [0, 2\pi[$. The activation gun is used to activate other agents. When another agent responds to a shot by the activation gun, the shooting agent switches to standby (only responding if hit). With two search guns an agent can particularly perform a move into direction $\beta \in [0, 2\pi[$ under
*Rule (∗):* Move until either the position of an other agent is reached or until there is an agent in the directions $\beta + \frac{\pi}{2}$ or $\beta - \frac{\pi}{2}$. (Move while constantly shooting left and right with search gun until someone is hit.)
Initially the $n$ agents are located at different positions in space and the multi-agent protocol is started form the outside by activating one agent. Whenever an agent is activated it executes the following program:

search gun all around shot, detect $A$ as set of all directions where agents are
select $\alpha, \gamma$ such that for all $-1 \leq c \leq 1$ it holds $(\alpha + c\gamma) \mod 2\pi \notin A$ and $\gamma$ maximal
**if** $\gamma \geq \frac{\pi}{2} + \frac{\pi}{n}$ **then**
  tie agents at same position to move together
  move direction $\beta = \alpha + 2\pi \mod 2\pi$ with rule (∗)
**end if**
activation gun all around shot (random start) until someone hit
**if** no one hit **then**
  give signal 'consensus found!'
**end if**

The protocol ensures that always only one agent is activated when an agent finishes its action unless consensus is found. It also always leads to the movement of an agent after some time unless consensus is found, because for every configuration there is always at least one agent whose position is an extreme point of the convex hull such that the exterior angle of the convex hull is larger than $\pi + \frac{2\pi}{n}$ and thus $\gamma \geq \frac{\pi}{2} + \frac{\pi}{n}$. This is because $\pi + \frac{2\pi}{n}$ is the exterior angle of a regular polygon with $n$ edges, which is the 'worst case'-polygon. It is 'worst case', because it has from all polygons with $n$ edges the largest minimal exterior angle. Thus, the random search for an agent which finds a direction $\alpha$ always ends successfully unless consensus is reached. So, the protocol leads to a series of actions which either continues forever including movements forever or finishing when consensus is reached. We group actions to form a series of updating maps $f_t$. We group by the following rule: Starting with the first action we collect actions in the same group until an agent is found which moves. The next updating map $f_2$ is formed analog starting with the next action, and so on. Thus we have a series of update maps.

It is simple to see that the series of updating maps $f_1, f_2, \ldots$

[1] Here the matrix $\mathcal{A}(t)$ is the arbitrarily chosen communication topology at time $t$ and $x$ is a given state. The set $e_k(\mathcal{A}(t))(x)$ is a subset of the relative interior of the convex hull of the neighbors of $k$ (including $k$) in the current communication topology, and it determines the set where the state of node $k$ has to remain in after one iteration. So, $e_k$ has to be fixed for a given communication topology and a certain state regardless of the chosen updating map $f(t, \cdots)$. This is in analogy to equiproper which implies the existence of a minimal Hausdorff distance $\delta(x)$ after one iteration for a given state but all possible averaging maps.

fulfills the conditions of Theorem 2.4 with $y$ the identity. Every $f_t$ is an averaging map because by definition the movement of agents goes into the convex hull or along its border and stops before the convex hull is left. It is equiproper, because for each $x$ there are only as many possible updating maps as their convex hull has extreme points. Thus, there is a $\delta(x) > 0$ by taking the minimum over this finite set of possible updating maps. Every $f_t$ is continuous in $x$ when we regard all agents which have the same position as one agent. Equicontinuity at $x$ again follows from finiteness of the possible updating maps.

Thus, the protocol in Example 4.4 leads to convergence to consensus. This can not be shown by applying Moreaus's Theorem 2 because the movements cannot be easily encoded in terms of communication topologies. One could try to specify it in terms of communication topologies by stating that the moving agent has agents at the detected directions in $A$ as its set of neighbors. But even then the conditions of Moreau's Assumption 1 (especialy number 3) need not be fulfilled and a node connected to all other nodes across time intervals of length $T$ need not exist as necessary for Moreau's Theorem 2.

*Example 4.5 (Nonlinear proper averaging map):* Let

$$f_1(x) = x_1,\ f_2(x) = a(l)x_1 + (1 - a(l))x_2,\ f_3(x) = \tfrac{1}{5}x_2 + \tfrac{4}{5}x_3$$

where $l = \mathrm{dist}(x_3, \text{line passing through } x_1 \text{ and } x_2)$ and $a$ is continuous and decreasing from $\frac{1}{2}$ to 0 in $[0, 1]$ and zero otherwise. In this example agent 3 moves towards agent 2 while agent 2 moves towards agent 1 only if agent 3 is close to a stripe around the line through agent 2 and agent 1.

Examples of this kind can be formulated in terms of communication topologies as Moreau's Theorem 2 needs them, but the existance of a uniform bound for the length of intercommunication intervals $T$ is not easily at hand.

*Example 4.6 (Non-arithmetic means):* We define $g_1, g_2, g_3, g_4 : (\mathbb{R}^d)^3 \to \mathbb{R}^d$ by $g_1(x) := \max\{x^1, x^2, x^3\}$, $g_2(x) := \frac{1}{3}(x^1 + x^2 + x^3)$, $g_3(x) := \sqrt[3]{x^1x^2x^3}$ and $g_4(x) := \min\{x^1, x^2, x^3\}$ with all computations componentwise. Further on let $f^{\sigma_1\sigma_2\sigma_3} : (\mathbb{R}^d)^3 \to (\mathbb{R}^d)^3$ with

$$f^{\sigma_1\sigma_2\sigma_3} := (g_{\sigma_1}, g_{\sigma_2}, g_{\sigma_3}).$$

It is easy to verify, that the family of all $f^{\sigma_1\sigma_2\sigma_3}$ where 1 and 4 are not both in $(\sigma_1, \sigma_2, \sigma_3) \in \{1, 2, 3, 4\}^3$ is an equicontinuous set of $y$-averaging maps, when the $y$-convex hull is the interval $[\min_{i\in\underline{n}} x^i, \max_{i\in\underline{n}} x^i]$. Equiproper is implied by finiteness. Thus convergence to consensus is ensured by Theorem 2.4. Moreau's theorem is not applicable because $f^{\sigma_1\sigma_2\sigma_3}$ is not a convex hull averaging map if some $\sigma_i$ is 1, 3 or 4 (since the componentwise $\min$ or $\max$ and the geometric mean are in general not contained in the convex hull).

Krause [10] shows another example where Moreau's theorem does not imply convergence: Assume three agents in two dimensional space. In each iteration every agent takes the mean value of the two other agents. Hence, no agent moves into the relative interior of the convex hull but these maps are still proper averaging maps and Theorem 2.4 applies.

## APPENDIX A
## PROOFS

*Proof of Lemma 2.3:* First we show that $g$ is an averaging map. Take $x \in S^n$ and let $\varepsilon > 0$. Due to the pointwise convergence of $(f_t)_i$ to $g_i$ and uniform continuity of $y$ there is $t_0$ such that for all $t > t_0$ it holds $\|y^i(f_t(x)) - y^i(g(x))\| < \varepsilon$. Due to $y^i(f_t(x)) \in \mathrm{conv}_{i\in\underline{m}}\, y^i(x)$ it follows that the maximal distance of $y^i(g(x))$ to $\mathrm{conv}_{i\in\underline{m}}\, y^i(x)$ is less than $\varepsilon$, and thus $y^i(g(x)) \in \mathrm{conv}_{i\in\underline{m}}\, y^i(x)$ because $\mathrm{conv}_{i\in\underline{m}}\, y^i(x)$ is closed.

We show that $g$ is proper. To this end, let $x \in S^n$ be not a consensus. We have to show that there is $z^* \in \mathrm{conv}_{i\in\underline{m}}\, y^i(x)$ but $z^* \notin \mathrm{conv}_{i\in\underline{m}}\, y^i(g(x))$. (Note that $z^* \in S$, while $x \in S^n$ and $y(x) \in S^m$.) We know that there is for each $t \in \mathbb{N}$ an $z(t) \in \mathrm{conv}_{i\in\underline{m}}\, y^i(x)$ with $z(t) \notin \mathrm{conv}_{i\in\underline{m}}\, y^i(f_t(x))$. According to the equiproper property it can be chosen such that the distance of $z(t)$ to $\mathrm{conv}_{i\in\underline{m}}\, y^i(f_t(x))$ is bigger than $\frac{\delta(x)}{2} > 0$ for all $t \in \mathbb{N}$. Further on, we know that the set difference $\mathrm{conv}_{i\in\underline{m}}\, y^i(f_t(x)) \setminus \mathrm{conv}_{i\in\underline{m}}\, y^i(x)$ is non empty and bounded, thus there is a subsequence $t_s$ such that $z(t_s)$ converges to a $z^* \in \mathrm{conv}_{i\in\underline{m}}\, y^i(x)$. Because of the construction it also holds $z^* \notin \mathrm{conv}_{i\in\underline{m}}\, y^i(g(x))$. ■

*Proof of Theorem 2.4:* The idea of the proof is the following: We define $C(t) := \mathrm{conv}_{i\in\underline{m}}\, y^i(x(t))$ which is convex and compact. It holds $C(t + 1) \subset C(t)$ because of the averaging property and $C := \bigcap_{t=0}^{\infty} C(t) \neq \emptyset$ because of compactness. In the following we will show that $C$ is a singleton, and that for all $i \in \underline{n}$ the sequences $x^i(t)$ converge to it. This will be done in three main steps, but first we note that because of compactness of $C(0)^n$ there is a subsequence $t_s$ and $c := (c^1, \ldots, c^n) \in C(0)^n$ such that $\lim_{s\to\infty} x(t_s) = c$.

1) We show that $C = \mathrm{conv}_{i\in\underline{m}}\, y^i(c)$. To accept "$\supset$" see that for all $t_s \geq t$ there is $x^i(t_s) \in C(t)$ and thus $c^i \in C(t)$. This implies $c^i \in C$ because all the $C(t)$ are closed.
   To show "$\subset$" let $x \in C$ and $\varepsilon > 0$. Because of uniform continuity of $y$ there is $\eta > 0$ such that for every $x' \in S$ with $\|c - x'\| < \eta$ it holds $\|y(c) - y(x')\| < \varepsilon$. Further on, there is $s_0$ such that for all $s \geq s_0$ it holds $\|x(t_s) - c\| < \eta$. This implies for every $i \in \underline{m}$ that $\left\|y^i(x(t_s)) - y^i(c)\right\| < \varepsilon$.
   Obviously, $x \in C(t_{s_0})$. Thus, there exist convex coefficients $a_1, \ldots, a_m \in \mathbb{R}^d_{\geq 0}$ such that $x = \sum_{i=1}^m a_i y^i(x(t_{s_0}))$. Now we can conclude
$$\begin{aligned}\|x - \sum_{i=1}^m a_i y^i(c)\| &= \|\sum_{i=1}^m a_i(y^i(x(t_{s_0})) - y^i(c))\| \\ &\leq \sum_{i=1}^m \|y^i(x(t_{s_0})) - y^i(c)\| = m\varepsilon.\end{aligned}$$
   It follows that $x \in \mathrm{conv}_{i\in\underline{m}}\, y^i(c)$ because $\mathrm{conv}_{i\in\underline{m}}\, y^i(c)$ is closed.
2) The next step is to show that $c$ is a consensus, i.e. $c^1 = \cdots = c^n$. The family $F$ is uniformly equicontinuous and for all $x \in X$ it holds that $\{f(x) \mid f \in F\}$ is bounded (and thus relatively compact) because all the $f$ are averaging maps. So, due to the theorem of Arzelà-Ascoli, $F$ is relatively compact. Thus, there is a subsequence $t_{s_r}$ such that $f_{t_{s_r}}$ converges uniformly to a continuous limit function $g$ for $r \to \infty$. Due to Lemma 2.3 we also know that $g$ is a proper averaging map. In two substeps we show that $\mathrm{conv}_{i\in\underline{m}}\, y^i(g(c)) = \mathrm{conv}_{i\in\underline{m}}\, y^i(c)$ which implies that $c$ is a consensus:
   a) We show that for all $i \in \underline{n}$ it holds $\lim_{r\to\infty} f_{t_{s_r}}(x_{t_{s_r}}) = g(c)$. We know that $f_{t_{s_r}} \to g$ uniformly and that $x^i(t_{s_r}) \to c$. Now we estimate
$$\begin{aligned}\|f_{t_{s_r}}(x(t_{s_r})) - g(c)\| \leq &\|f_{t_{s_r}}(x(t_{s_r})) - f_{t_{s_r}}(c)\| \\ &+ \|f_{t_{s_r}}(c) - g(c)\|\end{aligned}$$
      Both terms on the right hand side can be smaller than $\frac{\varepsilon}{2}$ for any $\varepsilon$ for large enough $r$ because of the continuity of $f_{t_{s_r}}$ and the uniform convergence $f_{t_{s_r}} \to g$.
   b) We show $\mathrm{conv}_{i\in\underline{m}}\, y^i(g(c)) = \mathrm{conv}_{i\in\underline{m}}\, y^i(c)$. "$\subset$" holds because $g$ is an averaging map (see 2a). To show "$\supset$" let $x \in \mathrm{conv}_{i\in\underline{m}}\, y^i(c)$. Thus, for all $r$ it holds $x \in C \subset C(t_{s_r} + 1)$ and thus there exist convex coefficients with convex combination $x = \sum_{i=1}^m a_i(r) y^i(x(t_{s_r} + 1))$.

Now, $(a_1(r),\dots,a_m(r))_{r\in\mathbb{N}}$ is a sequence in the compact set of convex coefficients and thus there is a subsequence $r_q$ such that $\lim_{q\to\infty}(a_1(r_q),\dots,a_m(r_q)) = (a_1{}^*\dots a_m{}^*)$. Now due to 2c and continuity of $y$ it holds,

$$x = \sum_{i=1}^{m} \lim_{q\to\infty} a_i(r_q) \lim_{q\to\infty} y^i(x(t_{s_{r_q}}+1)) = \sum_{i=1}^{m} a_i{}^* y^i(g(c)).$$

Thus, $x \in \operatorname{conv}_{i\in\underline{m}} y^i(g(c))$.

This implies that $c$ is a consensus, because $g$ is a proper averaging map.

3) Finally, we show that for each $i\in\underline{n}$ the sequence $(x^i(t))_{t\in\mathbb{N}}$ (and not only subsequences) converges to $\gamma := c^1 = \dots = c^n$ for $t\to\infty$. We know that for $\varepsilon > 0$ there is a $r_0$ such that for each $i\in\underline{n}$ it holds $\|y^i(x(t_{s_{r_0}})) - \gamma\| < \varepsilon$. Further on, for $t \geq t_{s_{r_0}}$ it holds $x(t) \in C(t) \subset C(t_{s_{r_0}})$. Thus, for each $i\in\underline{n}$ there are convex combinations $x^i(t) = \sum_{j=1}^{m} a^j y^j(x(t_{s_{r_0}}))$. Now, we conclude for all $t > t_{s_{r_0}}$

$$\|x^i(t) - \gamma\| = \|\sum_{j=1}^{n} a^j y^j(x(t_{s_{r_0}})) - \gamma)\| \leq \sum_{j=1}^{m} \|x^j(t_{s_{r_0}}) - \gamma\| = m\varepsilon.$$

This proves the theorem. ∎

*Proof of Lemma 2.5:* Let $\varepsilon > 0$ and consider two initial values $x(0), \tilde{x}(0) \in S^n$ with corresponding limits $\gamma$, $\tilde{\gamma}$ respectively. We have to show that there exists $\delta > 0$ such that $d(x(0),\tilde{x}(0)) \leq \delta$ implies $d(\gamma,\tilde{\gamma}) \leq \varepsilon$.

We note that for every $t$ it holds that

$$d(\gamma,\tilde{\gamma}) \leq d(\gamma, x(t)) + d(x(t),\tilde{x}(t)) + d(\tilde{\gamma},\tilde{x}(t)).$$

We choose $t_0$ large enough, that

$$d(\gamma, x(t_0)) \leq \frac{\varepsilon}{3} \qquad d(\tilde{\gamma},\tilde{x}(t)) \leq \frac{\varepsilon}{3}.$$

Since $\{f_t\}$ is an equicontinuous family there exists $\eta > 0$ such that for every $t\in\mathbb{N}$ it holds that

$$d(x(t),\tilde{x}(t)) \leq \eta \implies d(f_t(x(t)), f_t(\tilde{x}(t))) \leq \varepsilon.$$

Since $x(t)$ and $\tilde{x}(t)$ solve $x(t+1) = f_t(x(t))$ we have recursively that for every $t_0$ there exists $\delta > 0$ such that

$$d(x(0),\tilde{x}(0)) \leq \delta \implies d(x(t_0),\tilde{x}(t_0)) \leq \frac{\varepsilon}{3}$$

which implies the claim. ∎

*Proof of Theorem 2.7:* We define $f_t = \phi\circ g_t\circ\phi^{-1} : S^n \to S^n$. We show that $\{f_t \mid t\in\mathbb{N}\}$ is a family of equicontinuous, equiproper $y$-averaging maps on $S^n$. Equicontinuity and the fact that the $f_t$'s are $y$-averaging maps are clear. To see equiproperness of $f_t$ we note first that from equiproperness of $g_t$ it follows

$$d_H\big(\phi^{-1}(\operatorname*{conv}_{i\in\underline{m}} y^i(\phi(g_t(x)))), \phi^{-1}(\operatorname*{conv}_{i\in\underline{m}} y^i(\phi(x)))\big) \geq \delta(x)$$
$$\implies d_H(\phi^{-1}(\operatorname*{conv}_{i\in\underline{m}} y^i(f_t(\xi))), \phi^{-1}(\operatorname*{conv}_{i\in\underline{m}} y^i(\xi))) \geq \delta(\phi^{-1}(\xi))$$

while the second line holds for all $\xi = \phi(x) \in S^n$ and $t \geq 0$. Due to (3) we can express the Hausdorff distance as

$$\begin{aligned} & d_H(\phi^{-1}(\operatorname*{conv}_{i\in\underline{m}} y^i(f_t(\xi))), \phi^{-1}(\operatorname*{conv}_{i\in\underline{m}} y^i(\xi))) \\ &= \max_{z\in\phi^{-1}(\operatorname{conv}_{i\in\underline{m}} y^i(f_t(\xi)))} \; \min_{w\in\phi^{-1}(\operatorname{conv}_{i\in\underline{m}} y^i(\xi))} \|z-w\| \\ &= \max_{\phi(z)\in\operatorname{conv}_{i\in\underline{m}} y^i(f_t(\xi))} \; \min_{\phi(w)\in\operatorname{conv}_{i\in\underline{m}} y^i(\xi)} \|z-w\|. \end{aligned}$$

With this preparation we show equiproperness of the $f_t$'s:

$$\begin{aligned} & d_H\big(\operatorname*{conv}_{i\in\underline{m}} y^i(f_t(\xi)), \operatorname*{conv}_{i\in\underline{m}} y^i(\xi)\big) \\ &= \max_{\zeta\in\operatorname{conv}_{i\in\underline{m}} y^i(f_t(\xi))} \; \min_{\omega\in\operatorname{conv}_{i\in\underline{m}} y^i(\xi)} \|\zeta-\omega\| \\ &= \max_{\phi(z)\in\operatorname{conv}_{i\in\underline{m}} y^i(f_t(\xi))} \; \min_{\phi(w)\in\operatorname{conv}_{i\in\underline{m}} y^i(\xi)} \|\phi(z)-\phi(w)\| \\ &\geq L \max_{\phi(z)\in\operatorname{conv}_{i\in\underline{m}} y^i(f_t(\xi))} \; \min_{\phi(w)\in\operatorname{conv}_{i\in\underline{m}} y^i(\xi)} \|z-w\| \\ &\geq L\delta(\phi^{-1}(\xi)) \end{aligned}$$

where $L$ is the Lipschitz constant of $\phi^{-1}$. Now for $\xi(t) = \phi(x(t))$ it follows $\xi(t+1) = \phi(g_t(x(t)) = \phi(g_t(\phi^{-1}(\xi(t)))) = f_t(\xi(t))$. By virtue of Theorem 2.4, $\xi(t) \to c$ where $c\in S^n$ is a consensus and hence, $x(t)\to\phi^{-1}(c)\in T^n$ which is also a consensus. ∎

## References


[1] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. dissertation, Department of EECS, MIT, November 1984.

[2] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *Automatic Control, IEEE Transactions on*, vol. 31, no. 9, pp. 803–812, Sep 1986.

[3] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *Automatic Control, IEEE Transactions on*, vol. 48, no. 6, pp. 988–1001, June 2003.

[4] R. Saber and R. Murray, "Flocking with obstacle avoidance: cooperation with limited communication in mobile networks," vol. 2, Dec. 2003, pp. 2022–2028 Vol.2.

[5] V. Blondel, J. Hendrickx, A. Olshevsky, and J. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, pp. 2996–3000, Dec. 2005.

[6] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *Automatic Control, IEEE Transactions on*, vol. 50, no. 2, pp. 169–182, Feb. 2005.

[7] W. Ren, R. Beard, and E. Atkins, "A survey of consensus problems in multi-agent coordination," *American Control Conference, 2005. Proceedings of the 2005*, pp. 1859–1864 vol. 3, June 2005.

[8] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65 – 78, 2004.

[9] D. Angeli and P.-A. Bliman, "Stability of leaderless discrete-time multi-agent systems," *Mathematics of Control, Signals, and Systems (MCSS)*, vol. 18, no. 4, pp. 293–322, Oct. 2006.

[10] U. Krause, "Compromise, consensus, and the iteration of means," *Elemente der Mathematik*, vol. 64, no. 1, pp. 1–8, 2009.

[11] E. Seneta, *Non-Negative Matrices and Markov Chains*. Springer, 2006.

[12] J. Lorenz, "Repeated Averaging and Bounded Confidence-Modeling, Analysis and Simulation of Continuous Opinion Dynamics," Ph.D. dissertation, Universität Bremen, March 2007. [Online]. Available: http://nbn-resolving.de/urn:nbn:de:gbv:46-diss000106688

[13] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence, models, analysis and simulation," *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, p. 2, 2002. [Online]. Available: http://jasss.soc.surrey.ac.uk/5/3/2.html

[14] J. Lin, A. Morse, and B. Anderson, "The multi-agent rendezvous problem," *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, vol. 2, pp. 1508–1513 Vol.2, Dec. 2003.