Hegselmann-Krause Dynamics†

&

Constrained Consensus*

Angelia Nedić

angelia@illinois.edu

ISE Department and Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

joint work
with †B. Touri (former PhD student), and *P.A. Parrilo and A. Ozdaglar
Bio-Inspired (State-Dependent) Dynamic

- Bio-inspired models gaining popularity across vast areas of research
  - Dynamics and Control: flight formation, collision avoidance, coordination of network of robots
  - Signal Processing: distributed in network estimation and tracking
  - Optimization and Learning: distributed multi-agent optimization over network, distributed resource allocation, distributed learning, data-mining
  - Social/Economical Behavior: social networks, spread of influence, opinion formation and dynamics

- At the core of all these is a bio-inspired dynamical model for information spread over a network through local agents interactions
Multi-dimensional Hegselmann-Krause Model

- Discrete-time model used often for opinion dynamics and robotic networks
- The model consists of the following
  - The set \([m] = \{1, 2, \ldots, m\}\) of \(m\) agents
  - The agents interact at discrete time instances, indexed by \(t = 1, 2, 3, \ldots\)
  - Each agent \(i\) has an opinion modeled by a vector \(x_i(t) \in \mathbb{R}^n\) at time \(t\)
  - Starting with some initial opinion \(x_i(0) \in \mathbb{R}^n\), each agent updates his opinion upon interaction with neighbors
  - Neighbors are defined in terms of a confidence interval, given by a scalar \(\epsilon > 0\)
    \[\mathcal{N}_i(t) = \{j \in [m] \mid \|x_j(t) - x_i(t)\| \leq \epsilon\}\]
  - Each agent updates his opinion by averaging the opinions of all his neighbors
    \[x_i(t + 1) = \frac{1}{|\mathcal{N}_i(t)|} \sum_{j \in \mathcal{N}_i(t)} x_j(t) \quad \text{for } t = 0, 1, 2, \ldots\]

Proposed by Hegselmann and Krause 2002* for scalar case

What is known about Hegselmann-Krause Model

\[ x_i(t + 1) = \frac{1}{|\mathcal{N}_i(t)|} \sum_{j \in \mathcal{N}_i(t)} x_j(t) \quad \text{for } t = 0, 1, 2, \ldots \]

\[ \mathcal{N}_i(t) = \{ j \in [m] \mid \|x_j(t) - x_i(t)\| \leq \epsilon \} \quad \text{for } t = 0, 1, \ldots \]

SCALAR CASE

- Given $\epsilon > 0$ and the initial profile $\{x_i(0), i \in [m]\}$ the behavior of the opinions is completely determined
- For any initial profile and any confidence interval $\epsilon$, the agents opinions $x_i(t)$ converge to some limiting opinion $x_i^*$ in a finite time, i.e., for each agent $i$, there exists time $T_i \geq 0$ such that

  \[ x_i(t) = x_i^* \quad \text{for all } t \geq T_i \]

- The fragmentation of opinion may occur i.e., we may have agents $i$ and $j$ such that

  \[ x_i^* \neq x_j^* \]

- Hegselmann and Krause 2002, Blondel et al. 2009†
- Convergence time of the order $O(m^4)$ - Behrouz Touri Ph.D. Thesis
- Improved to $O(m^3)$ - S. Mohajer and B. Touri, On Convergence Rate of Scalar Hegselmann-Krause Dynamics, ACC 2013

---

Some Illustrations of the Opinion Dynamics

- 9 groups in the limit, convergence in 16 time steps
- The fragmentations of opinions occur, BUT the fragmentations are permanent
- The analysis of the scalar case depends heavily on the fact that the ordering of the agents is not changing, when the agents are ordered according to the values they have.
Illustration of the Opinion Dynamics in 2-D

- The opinions can merge!!!

- The fragmentations of opinions may occur, BUT the fragmentations are NOT permanent

- The analysis cannot make use of the ordering of the agents’ values
What is known about HK Model

- Convergence in a finite time - Lorenz 2007 (PhD Thesis)

- Not knowing this result at the time we have shown that the dynamics converges in a finite time for any initial profile and confidence $\epsilon > 0$

- Lorenz’s proof considers the matrix products - convergence rate - hard to get from that analysis

- Our proof is based on a dynamic system point of view and makes use of two basic ingredients
  - An adjoint dynamics - existence
  - Selection of a Lyapunov function - using adjoint variables
  - Open question - ???
Multidimensional Hegselmann-Krause (HK) Model

Agent Types

\[ M(t) = \{ i \in [m] \mid \max_{j \in N_i(t)} ||x_i(t) - x_j(t)|| > \epsilon / 2 \}, \]

\[ G(t) = \{ i \in [m] \mid \max_{j \in N_i(t)} ||x_i(t) - x_j(t)|| \leq \epsilon / 2 \}. \]

If an agent and all his neighbors have small local-opinion spread, then the agent and his neighbors will reach an agreement in the next time step. **Left** An agent and all his neighbors have small local-opinion spread, which leads to their agreement at next time step. **Right** An agent with a small local-opinion spread, with a neighbor whose local-opinion spread is large.
Lemma 1 There holds for all $t \geq 0$,
(a) $N_i(t) \subseteq N_j(t)$ for all $i \in G(t)$ and all $j \in N_i(t)$.

(b) $N_j(t) = N_i(t)$ for all $i \in G(t)$ and all $j \in N_i(t) \cap G(t)$.

(c) For every $i, \ell \in G(t)$, either $N_i(t) = N_\ell(t)$ or $N_i(t) \cap N_\ell(t) = \emptyset$.

A consequence of Lemma 1(b), for $i \in G(t)$ with $N_i(t) \subseteq G(t)$, we have
$$x_i(t + 1) = x_j(t + 1) \quad \text{for all } j \in N_i(t).$$
**Left:** an agent in $G(t)$ with a neighbor not in $G(t)$; **Right:** an agent in $G(t)$ with all neighbors in $G(t)$
**Important Relation**

**Lemma 2** Let $i \in G(t)$ be such that $\mathcal{N}_i(t) \subseteq G(t)$. Then, at time $t + 1$, one of the following two cases occurs

$$\mathcal{N}_i(t + 1) = \mathcal{N}_i(t) \quad \text{or} \quad \mathcal{N}_i(t + 1) \supset \mathcal{N}_i(t).$$

Furthermore, in the second case if there is an agent $\ell \in G(t)$ such that $\ell \in \mathcal{N}_i(t + 1) \setminus \mathcal{N}_i(t)$, then $\ell \in \mathcal{N}_j(t + 1)$ for all $j \in \mathcal{N}_i(t)$ and

$$\|x_j(t + 1) - x_\ell(t + 1)\| > \frac{\epsilon}{2} \quad \text{for all} \ j \in \mathcal{N}_i(t).$$
Termination Criteria

- The termination time $T$ of the HK dynamics is defined by

$$T = \inf_{t \geq 0} \{ t \mid x_i(k + 1) = x_i(k) \text{ for all } k \geq t \text{ and all } i \in [m] \},$$

**Proposition 1** Suppose that the time $\hat{t}$ is such that $G(\hat{t}) = [m]$ and $G(\hat{t} + 1) = [m]$.

Then, we have

$$N_i(\hat{t} + 1) = N_i(\hat{t}) \quad \text{for all } i \in [m],$$

and the termination time $T$ of the Hegselmann-Krause dynamics satisfies $T \leq \hat{t} + 1$.

Since $G(\hat{t}) = [m]$, by Lemma 2, it follows that either $N_i(\hat{t} + 1) = N_i(\hat{t})$ for all $i \in [m]$, or $N_i(\hat{t} + 1) \supset N_i(\hat{t})$ for some $i \in [m]$. We show that the latter case cannot occur. Specifically, every $\ell \in N_i(\hat{t} + 1) \setminus N_i(\hat{t})$ must belong to the set $G(\hat{t})$ since $M(\hat{t}) = \emptyset$. Therefore, by Lemma 2 it follows that $\|x_j(\hat{t} + 1) - x_\ell(\hat{t} + 1)\| > \frac{\epsilon}{2}$ for all $j \in N_i(\hat{t})$. In particular, this implies $i, \ell \in M(\hat{t} + 1)$ which is a contradiction since $M(\hat{t} + 1) = \emptyset$. Therefore, we must have $N_i(\hat{t} + 1) = N_i(\hat{t})$ for all $i$, which means that each agent group has reached a local group-agreement and this agreement persists at time $\hat{t} + 1$ and onward.
**Adjoint Dynamics**

We next show that the Hegsellman-Krause dynamics has an adjoint dynamics. To do so, we first compactly represent the dynamics by defining the neighbor-interaction matrix $B(t)$ with the entries given as follows:

$$B_{ij}(t) = \begin{cases} \frac{1}{|N_i(t)|} & \text{if } j \in N_i(t), \\ 0 & \text{otherwise}. \end{cases}$$

The HK dynamics can now be written as:

$$X(t + 1) = B(t)X(t) \quad \text{for all } t \geq 0,$$

where $X(t)$ is the $m \times n$ matrix with the rows given by the opinion vectors $x_1(t), \ldots, x_m(t)$.

- We say that the Hegselmann-Krause dynamics has an *adjoint dynamics* if there exists a sequence of probability vectors $\{\pi(t)\} \subset \mathbb{R}^m$ such that

  $$\pi'(t) = \pi'(t + 1)B(t) \quad \text{for all } t \geq 0.$$

- We say that an adjoint dynamics is uniformly bounded away from zero if there exists a scalar $p^* \in (0, 1)$ such that $\pi_i(t) \geq p^*$ for all $i \in [m]$ and $t \geq 0$. 
Existence of Adjoint Dynamics for HK model

To show the existence of the adjoint dynamics, we use a variant Theorem 4.8 in Behrouz Thesis, as stated below.

**Theorem 2** Let \( \{A(t)\} \) be a sequence of stochastic matrices such that the following two properties hold:

(i) There exists a scalar \( \gamma \in (0, 1) \) such that \( A_{ii}(t) \geq \gamma \) for all \( i \in [m] \) and \( t \geq 0 \);

(ii) There exists a scalar \( \alpha \in (0, 1] \) such that for every nonempty \( S \subset [m] \) and its complement \( \bar{S} = [m] \setminus S \), there holds

\[
\sum_{i \in S, j \in \bar{S}} A_{ij}(t) \geq \alpha \sum_{j \in \bar{S}, i \in S} A_{ji}(t) \quad \text{for all } t \geq 0.
\]

Then, the dynamics \( z(t + 1) = A(t)z(t), \ t \geq 0 \), has an adjoint dynamics \( \{\pi(t)\} \) which is uniformly bounded away from zero.

- HK dynamics is driven by stochastic matrices \( B(t) \) to which the Theorem applies

- **HK opinion dynamics has an adjoint dynamics which is uniformly bounded away from zero.**
Lyapunov Comparison Function

We construct a Lyapunov comparison function by using the adjoint HK dynamics. The comparison function for the dynamics is a function $V(t)$, which is defined for the $m \times n$ opinion matrix $X(t)$ and $t \geq 0$:

$$V(t) = \sum_{i=1}^{m} \pi_i(t) \|x_i(t) - \pi'(t)X(t)\|^2,$$

(3)

where the row $X_{i,:}(t)$ is given by the opinion vector $x_i(t)$ of agent $i$ and $\pi(t)$ is the adjoint dynamics. For this function, we have following essential relation.

**Lemma 3** For any $t \geq 0$, we have

$$V(t + 1) = V(t) - D(t)$$

where

$$D(t) = \frac{1}{2} \sum_{\ell=1}^{m} \frac{\pi_{\ell}(t + 1)}{|\mathcal{N}_{\ell}(t)|^2} \sum_{i \in \mathcal{N}_{\ell}(t)} \sum_{j \in \mathcal{N}_{\ell}(t)} \|x_i(t) - x_j(t)\|^2.$$

• Proof of this result is in Behrouz Touri’s thesis‡

Finite Termination Time

Proposition 3 The Hegselmann-Krause opinion dynamics (2) reaches its steady state in a finite time.

By summing the relation of essential Lemma for $t = 0, 1, \ldots, \tau - 1$ for some $\tau \geq 1$, and by rearranging the terms, we obtain

$$V(\tau) + \frac{1}{2} \sum_{t=0}^{\tau-1} \sum_{\ell=1}^{m} \frac{\pi_\ell(t + 1)}{|N_\ell(t)|^2} \sum_{i \in N_\ell(t)} \sum_{j \in N_\ell(t)} \|x_i(t) - x_j(t)\|^2 = V(0).$$

Letting $\tau \to \infty$, since $V(\tau) \geq 0$ for any $\tau$, we conclude that

$$\lim_{t \to \infty} \sum_{\ell=1}^{m} \frac{\pi_\ell(t + 1)}{|N_\ell(t)|^2} \sum_{i \in N_\ell(t)} \sum_{j \in N_\ell(t)} \|x_i(t) - x_j(t)\|^2 = 0.$$

By Theorem 2 (existence of adjoint dynamics) the adjoint dynamics $\{\pi(t)\}$ is uniformly bounded away from zero, i.e., $\pi_\ell(t) \geq p^*$ for some $p^* > 0$, and for all $\ell \in [m]$ and $t \geq 0$. Furthermore, $|N_\ell(t)| \leq m$ for all $\ell \in [m]$ and $t \geq 0$, Therefore, it follows that for every $\ell \in [m]$,

$$\lim_{t \to \infty} \sum_{i \in N_\ell(t)} \sum_{j \in N_\ell(t)} \|x_i(t) - x_j(t)\|^2 = 0.$$
We further have
\[ \sum_{i \in N} \ell (t) \sum_{j \in N} \ell (t) \| x_i(t) - x_j(t) \|^2 \geq \sum_{i \in N} \ell (t) \| x_i(t) - x_\ell(t) \|^2 \geq \max_{i \in N} \ell (t) \| x_i(t) - x_\ell(t) \|^2, \]
where the first inequality follows from the fact $\ell \in N(t)$ for all $\ell \in [m]$ and $t \geq 0$. Consequently, we obtain for every $\ell \in [m]$,
\[ \lim_{t \to \infty} \max_{i \in N} \ell (t) \| x_i(t) - x_\ell(t) \|^2 = 0. \]
Hence, for every $\ell \in [m]$, there exists a time $t_\ell$ such that
\[ \max_{i \in N} \ell (t) \| x_i(t) - x_\ell(t) \|^2 \leq \frac{\epsilon}{2} \quad \text{for all } t \geq t_\ell. \]
In view of the definition of the set $M(t)$ in (1), the preceding relation implies that
\[ M(t) = \emptyset \quad \text{for all } t \geq \max_{\ell \in [m]} t_\ell. \]
By Proposition 1 (termination criteria), it follows that the termination time $T$ satisfies
\[ T \leq \max_{\ell \in [m]} t_\ell + 1. \]
- **Multidimensional HK dynamics converges in a finite time**
- Sadly the result is not new: Jan Lorenz 2007

---

\[ ^{\S}J. \text{Lorenz, "Repeated averaging and bounded confidence modeling, analysis and simulation of continuous opinion dynamics," Ph.D. Dissertation, Universität Bremen, 2007.} \]
Some open directions/questions

- Convergence time?

- Can we predict the number of groups in the limit?

- What about convergence of such a dynamics when each agent uses its own $\epsilon_i$?

  Partial answers in
Constrained Consensus

- Intro to Feasibility Problems
- Constrained Consensus viewed as Feasibility Problem
- Distributed Algorithm for Constrained Consensus
Constrained Consensus Problem

- Given a sequence \( \{G_t\} \) of directed and strongly connected graphs, \( G_t = ([m], E_t) \)

- Each agent \( i \) (node in the graph) has a constraint set \( X_i \subseteq \mathbb{R}^n \), nonempty closed and convex

- Construct a process (algorithm) that uses local agent interactions, so that the agents reach an agreement on some vector \( x \in \bigcap_{i=1}^{m} X_i \)

- Sets \( X_i \) are assumed to be simple for projections

- The intersection set \( X = \bigcap_{i=1}^{m} X_i \) is assumed to be nonempty
Characterization of a convex set $C$

- Given a closed convex set $C$, consider the distance function associated with $C$:
  \[ \text{dist}(x, C) = \|x - \Pi_C[x]\| \quad \text{for all } x \in \mathbb{R}^n \]
  where the norm is the Euclidean norm

- Note that:
  \[ x \in C \iff \|x - \Pi_C[x]\| = 0 \]

- Hence:
  \[ x \in C \iff \text{dist}(x, C) = 0 \]

- The distance function is not differentiable, but the squared distance function is
  \[ g(x) = \text{dist}^2(x, C), \quad \nabla g(x) = 2(x - \Pi_C[x]) \]

- Thus, we can use the description of $C$ in terms of the squared distance function of $C$:
  \[ x \in C \iff \frac{1}{2} \text{dist}^2(x, C) = 0 \]
Consensus Problem Reformulation

- The set $X_i$ can be described in terms of the distance function $\mathbb{I} x \mapsto \text{dist}(x, X_i)$

$$X_i = \left\{ x \in \mathbb{R}^n \mid \frac{1}{2} \text{dist}^2(x, X_i) = 0 \right\}$$

- Consider the problem

$$\text{minimize} \sum_{i=1}^{m} \frac{1}{2} \text{dist}^2(x, X_i) \quad \text{subject to} \quad x \in \mathbb{R}^n$$

- We see that

finding a point $x \in X$ is equivalent to solving the preceding problem

- Letting $f_i(x) = \frac{1}{2} \text{dist}^2(x, X_i)$, we are back to the problem of the form

$$\text{minimize} \sum_{i=1}^{m} f_i(x) \quad \text{subject to} \quad x \in \mathbb{R}^n$$

\[
\text{AN "Random Projection Algorithms for Convex Set Intersection Problems" CDC 2010}
\]
• Each function \( f_i(x) = \frac{1}{2} \text{dist}^2(x, X_i) \) is convex (partial minimization criterion)

• Each function is differentiable

\[
\nabla f_i(x) = x - \Pi_{X_i}[x]
\n\]

• Furthermore, the gradients \( \nabla f_i \) for each \( i \) are Lipschitz continuous

\[
\|\nabla f_i(x) - \nabla f_i(y)\| = \|x - \Pi_{X_i}[x] - (y - \Pi_{X_i}[y])\| \\
\leq \|x - y\| + \|\Pi_{X_i}[x] - \Pi_{X_i}[y]\| \\
\leq 2\|x - y\|
\]

where the last inequality follows from the non-expansiveness property of the projection

\[
\|\Pi_{X_i}[x] - \Pi_{X_i}[y]\| \leq \|x - y\|
\]
Constrained Consensus Problem: New Formulation

- Given a sequence \( \{G_t\} \) of directed and strongly connected graphs, \( G_t = ([m], \mathcal{E}_t) \)

- Solve the problem

  \[
  \begin{align*}
  \text{minimize} & \quad \sum_{i=1}^{m} f_i(x) \\
  \text{subject to} & \quad x \in \mathbb{R}^n
  \end{align*}
  \]

  with \( f_i(x) = \frac{1}{2} \text{dist}^2(x, X_i) \)

- Each agent has access to \( f_i \) only

- USEFUL properties
  - Each \( f_i \) is Lipschitz-gradient convex function, with Lipschitz constant \( L_i = 2 \)
  - When \( X \triangleq \bigcap_{i=1}^{m} X_i \neq \emptyset \), the set of solutions of the problem is the set \( X \)
  - When \( X \neq \emptyset \), all functions \( f_i \) have a set of common global minima, which is the set \( X \)

- Apply distributed algorithm
Distributed Algorithm

- Let $N_i(t) = \{ j \in [m] | (j, i) \in \mathcal{E}_i \}$
- Let each agent have an initial vector $x_i(0) \in X_i$
- Agent updates:
  - **Consensus-like step** to mix their own estimate with those received from neighbors
    $$w_i(t+1) = \sum_{j=1}^{m} a_{ij}(t) x_j(t) \quad (a_{ij}(t) = 0 \text{ when } j \notin N_i(t))$$
  - Followed by a **local gradient-based step**
    $$x_i(t+1) = w_i(t+1) - \alpha(t) \nabla f_i(w_i(t+1))$$
    where $\alpha(t) > 0$ is a stepwise
- Algorithm in this case assumes a special form:
  - Use: $\nabla f_i(x) = x - \Pi_{X_i}[x]$
    $$x_i(t+1) = w_i(t+1) - \alpha(t) (w_i(t+1) - \Pi_{X_i}[w_i(t+1)])$$
    $$x_i(t+1) = (1 - \alpha(t)) w_i(t+1) + \alpha(t) \Pi_{X_i}[w_i(t+1)]$$
Distributed Algorithm: Stepize

• Agents initialize with $x_i(0) \in X_i$ for $i \in [m]$  
• At each update time $t$, every agent performs two steps:

$$w_i(t + 1) = \sum_{j=1}^{m} a_{ij}(t)x_j(t)$$

$$x_i(t + 1) = (1 - \alpha(t))w_i(t + 1) + \alpha(t)\Pi_{X_i}[w_i(t + 1)]$$

where $a_{ij}(t) = 0$ when $j \notin N_i(t)$ and $\alpha(t) > 0$

• Stepsizes here can take advantage of the gradient Lipschitz continuity:

  step-size can be constant $\alpha(t) = \alpha$ for all $t$

• Algorithm will be convergent as long as $\alpha \leq \min_i \frac{2}{L_i}$ where $L_i$ is Lipschitz constant for function $f_i$

• Recall $L_i = 2$ for all $f_i$ in this case

• So we can use $\alpha(t) = \alpha$ for all $t$, with $0 < \alpha \leq 1$.

• We consider the version with $\alpha = 1$

• Done in: AN, A. Ozdaglar and A.P. Parrilo "Constrained Consensus and Optimization in Multi-Agent Networks" IEEE TAC, 2010
Distributed Algorithm for Constrained Consensus

- Agents initialize with $x_i(0) \in X_i$ for $i \in [m]$.
- At each update time $t$, every agent $i$ updates as follows:

$$w_i(t + 1) = \sum_{j=1}^{m} a_{ij}(t)x_j(t), \quad x_i(t + 1) = \Pi_{X_i}[w_i(t + 1)]$$

- Compactly, agent updates have the form:

$$x_i(t + 1) = \Pi_{X_i} \left[ \sum_{j=1}^{m} a_{ij}(t)x_j(t) \right] \quad \text{for all } i \in [m] \text{ and all } t$$

Proposition (Convergence)\textsuperscript{II} Let each of the graphs $G_t$ be strongly connected and let each $A(t)$ be doubly stochastic matrix such that

$$a_{ii}(t) \geq \eta \quad \text{for all } i \text{ and } t$$
$$a_{ij}(t) \geq \eta \quad \text{for all } j \in N_i(t), \text{ and all } i \text{ and } t$$

where $\eta > 0$. Assume also that $X = \cap_{i=1}^{m} X_i \neq \emptyset$. Then, the iterate sequences $\{x_i(t)\}$, $i \in [m]$, converge to a common point $x^* \in X$, i.e.,

$$\lim_{t \to \infty} x_i(t) = x^* \quad \text{for some } x^* \in X \text{ and all } i \in [m]$$

\textsuperscript{II}Prop. 2 in AN, A. Ozdaglar and A.P. Parrilo, 2010
Convergence - open items

Proposition (Convergence)** Let each of the graphs $G_t$ be strongly connected and let each $A(t)$ be doubly stochastic matrix such that

$$a_{ii}(t) \geq \eta \quad \text{for all } i \text{ and } t$$

$$a_{ij}(t) \geq \eta \quad \text{for all } j \in N_i(t), \text{ and all } i \text{ and } t$$

where $\eta > 0$. Assume also that $X = \cap_{i=1}^{m} X_i \neq \emptyset$. Then, the iterate sequences $\{x_i(t)\}, i \in [m]$, converge to a common point $x^* \in X$, i.e.,

$$\lim_{t \to \infty} x_i(t) = x^* \quad \text{for some } x^* \in X \text{ and all } i \in [m]$$

- $B$-connectivity works (we have done that)

- Since all functions $f_i$ have a set of common minima when $X \neq \emptyset$, the assumption of the doubly stochastic matrices is NOT NEEDED

- It is sufficient to have row-stochastic matrices $A(t)$

- Convergence for this case is not done yet

**Prop. 2 in AN, A. Ozdaglar and A.P. Parrilo, 2010
Convergence Rate: Wide Open Question

Proposition (Convergence Rate)†† Let \( X = \bigcap_{i=1}^{m} X_i \neq \emptyset \) have an interior point, i.e., there is a vector \( \hat{x} \in X \) and a scalar \( \delta > 0 \) such that
\[
\{ z \in \mathbb{R}^n \mid \| z - \hat{x} \| \leq \delta \} \subseteq X
\]
Assuming that each graph \( G_t \) is complete and the weights \( a_{ij}(t) \) are uniform, i.e., \( A(t) = \frac{1}{m} \mathbf{1} \mathbf{1}' \), with \( \mathbf{1} \) being the vector of 1’s, we have
\[
\sum_{i=1}^{m} \| x_i(t) - x^* \|^2 \leq q t \sum_{i=1}^{m} \| x_i(0) - x^* \|^2 \quad \text{for all} \ t,
\]
where
\[
q = 1 - \frac{1}{4R^2}, \quad R = \frac{1}{\delta} \sum_{i=1}^{m} \| x_i(0) - x^* \|.
\]

• TOO STRONG assumptions
• The rate should be done under assumptions needed for convergence (no use doubly stochastic matrices) and some assumption for the set regularity
• Set regularity: interior point assumption is fine, but it is not the only one

††Prop. 3 in AN, A. Ozdaglar and A.P. Parrilo, 2010