DISTRIBUTED OPTIMIZATION (CONT.)

DISTRIBUTED DUAL AVERAGING:

for ea. node i.

AP for descent direction

$$\underline{z_i}(t+1) = \sum_{j \in N(i)} P_{ij}\, \underline{z_j(t)} - \underbrace{g_i(t)}_{\text{gradient of } f_i(x)} \qquad \underline{x_i}(t+1) = \Pi_{\mathcal{X}}^{\psi}\left(\underline{-z_i(t+1)}, \underline{\alpha(t)}\right)$$

$P \in \mathbb{R}^{n \times n}$ doubly stochastic matrix (graph structure)

scalar case. $z(t) = \begin{bmatrix} z_1(t) \\ \vdots \\ z_n(t) \end{bmatrix}$ $\quad z(t+1) = \underline{Pz(t)} - g(t) \quad g(t) = \begin{bmatrix} g_1(t) \\ \vdots \\ g_n(t) \end{bmatrix}$

$$\frac{1}{n}\mathbb{1}^T z(t+1) = \frac{1}{n}\mathbb{1}^T P z(t) - \frac{1}{n}\mathbb{1}^T g(t)$$

Average (over agents) descent direction

dynamics: $\bar{z}(t+1) = \bar{z}(t) - \frac{1}{n}\sum_{j=1}^{\hat{n}} \underline{g_j(t)}$

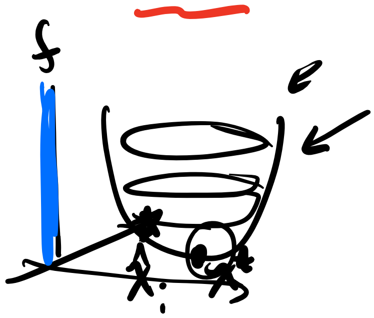
$$g_j = \left.\frac{\partial f_j}{\partial x}\right|_{x_j}$$

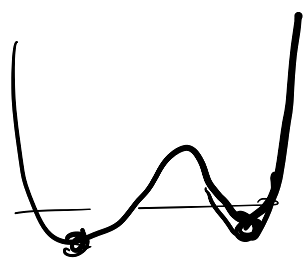
Thm 1: Basic Convergence

$$\left[\underbrace{f(\hat{x}_i(T))}_{\substack{\text{time average} \\ \text{of } x_i}} - \underbrace{f(x^*)}_{\text{optimizer}}\right] \le \frac{1}{T\alpha(T)}\psi(x^*) + \frac{L^2}{2T}\sum_{t=1}^{T}\alpha(t-1) + \boxed{\frac{3L}{T}\max_{j}\sum_{t=1}^{T}\alpha(t)\left\|\bar{z}(t) - z_j(t)\right\|_*}$$

usual terms
- Lipschitz f_i, ψ, $\Pi_{\mathcal{X}}^{\psi}$
- convexity of f_i
- g_i are gradients

NEW TERM





$\alpha(t) \sim \frac{1}{\sqrt{t}}$

as $\sigma_2 \to 1$

Thm 2: Rate of Convergence:

$\frac{1}{\sqrt{1-\sigma_2}}$ blows up

$\Rightarrow$ convergence is slower.

$$\alpha(t) = \underbrace{\frac{R\sqrt{1-\sigma_2(P)}}{4L}}_{\text{const.}}\left(\frac{1}{\sqrt{t}}\right)$$

new term.

$$f(\hat{x}_i(T)) - f(x^*) \leq 8\frac{RL}{\sqrt{T}}\frac{\log(T\sqrt{n})}{\boxed{\sqrt{1-\sigma_2(P)}}}$$

$\sigma_2(P)$ second largest singular value of P.

$1-\sigma_2(P)$: spectral gap.

P: doubly stochastic $\mathbb{1}^T P = \mathbb{1}^T$
symmetric $P\mathbb{1} = \mathbb{1}$.

P: eigenvalues

$$\underbrace{\lambda_1(P)}_{1} \geq \lambda_2(P) \geq \cdots \geq$$

since P is symmetric. singulars are magnitudes of eigenvalues

singular values $\underbrace{\sigma_1(P)}_{1} > \sigma_2(P) \geq \cdots \geq \sigma_n(P) \geq 0$

full connected graph.

spectral gap: $\sigma_1 - \sigma_2 = 1 - \sigma_2$

$\lambda_1 = \sigma_1 = 1$: Correspond to left eigenvector $\mathbb{1}^T$
right eigenvector $\underbrace{\mathbb{1}}_{\text{constant}}$

vectors stay the same

For different types of graphs...
& different weights in P.

$\sigma_2(P)$ different

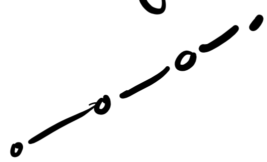
take P = related to graph Laplacian

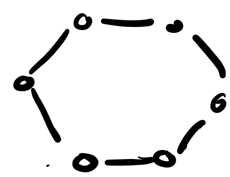
$\Rightarrow$ $\sigma_2(P)$

$\Rightarrow$ get convergence rate bounds.

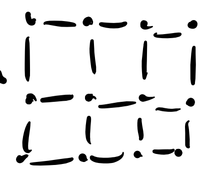
<u>Corollary 1</u>: apply $\sigma_2(P)$ results to get convergence rates for

(a) k-connected paths & cycles





(b) k-connected $\sqrt{n} \times \sqrt{n}$ grids
"a grid w/ n nodes"



(c) Random geometric graphs w/ connectivity radius r

(d) Expanders w/ bounded ratio of minimum to maximum node degree.

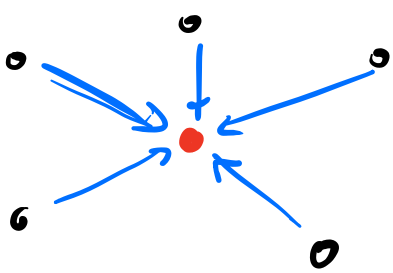
PROOF OF CONVERGENCE:

Summary "computing the settling time for the graph."

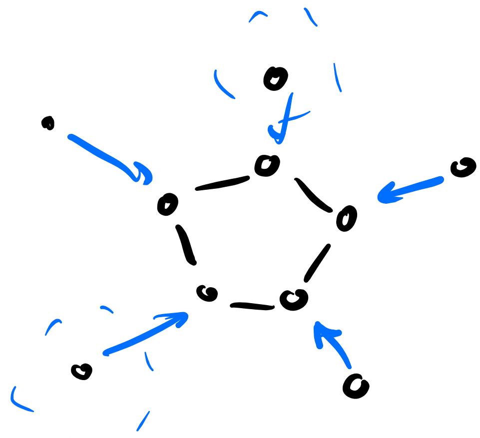
FORMATION CONTROL:

BEFORE

AP $x_i \in \mathbb{R}^2$ or $\mathbb{R}^3$



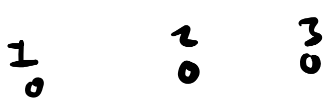
"Rendevous"

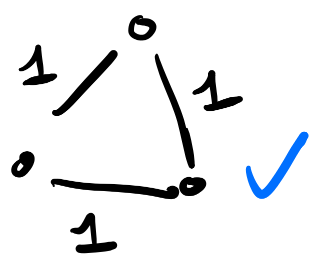


"Formation control"

Specifying a Formation

Relative distances $D = \{ d_{ij} \in \mathbb{R} \mid d_{ij} > 0 \quad i,j \in V \quad i \neq j \}$

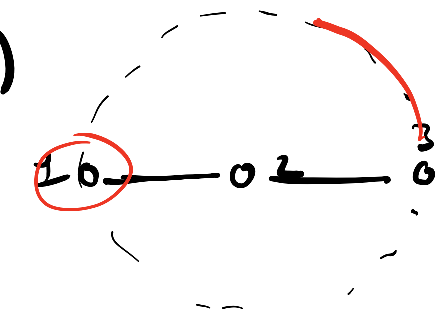




$D = \{ d_{12} = 1, d_{23} = 1, d_{13} = 1 \}$ good ✓

$D = \{ d_{12} = 1, d_{23} = 1, d_{13} = 3 \}$

$d_{13} < 2$

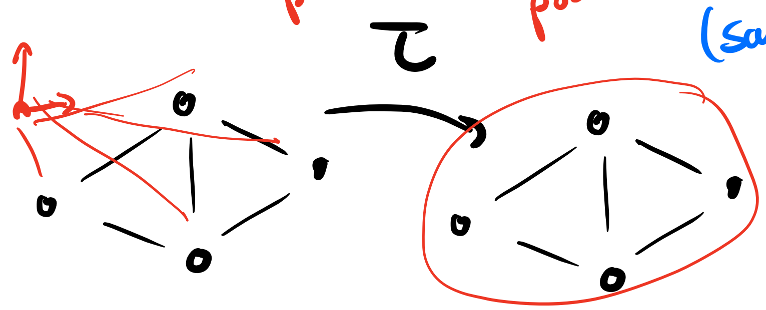


Relative Locations $\quad \Xi = \{ \xi_1, \dots, \xi_n \} \quad \xi_i \in \mathbb{R}^p \quad p = 2, 3$

relative positions

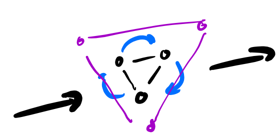
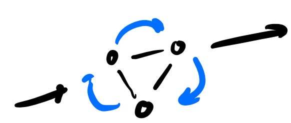
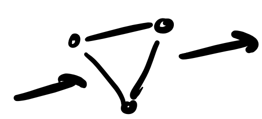
actual positions $\quad x_i = \xi_i + \boxed{\tau}$ group translation

actual position $\quad$ relative position $\quad$ arbitrary translation (same for all agents)



Types

more specific

formation	specification	interpretation	picture
scale invariant	D	$\lVert x_i - x_j \rVert = \alpha d_{ij}$ for $\alpha > 0$	
rigid	D	$\lVert x_i - x_j \rVert = d_{ij}$	
translation invariant	Ξ	$x_i = \xi_i + \tau$ for $\tau \in \mathbb{R}^p$	

Translation Invariant

easiest one to use
least pathologies (linear)

want to use AP
to do this $\rightarrow$ how?

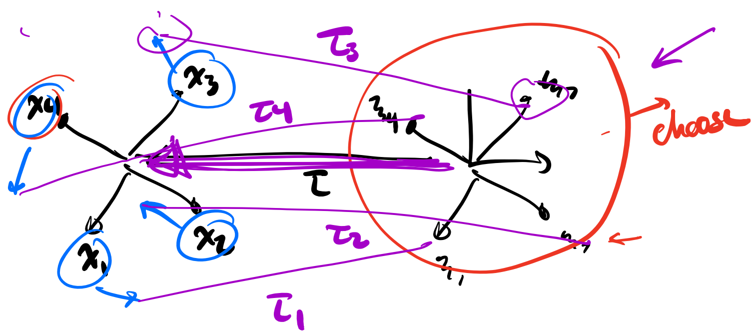
want to agree on common translation vector τ

ea agent keeps estimate of $\tau \rightarrow \tau_i$
AP.

$$\dot{\tau}_i(t) = \sum_{j \in N(i)} (\tau_i(t) - \tau_j(t)) \qquad \text{AP}$$

communication graph

$$x_i(t) = \xi_i + \tau_i(t) \Longrightarrow \tau_i(t) = x_i(t) - \xi_i$$



$$(x_i(t) - \cancel{\xi_i})^{\boldsymbol{\cdot}} = -\sum_{j \in N_i} \left[(x_i(t) - \xi_i) - (x_j(t) - \xi_j) \right]$$

$$\dot{x}_i(t) = -\sum_{j \in N_i} \Big[\underbrace{(x_i(t) - x_j(t))}_{} - \underbrace{(\xi_i - \xi_j)}_{\text{desired relative position}} \Big]$$

direction of motion for agent i

current position

neighbors position

desired relative position

in Matrix form: $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ $\xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$

Formation control $\equiv$ Rendezvous control

$$\dot{x} = -\underbrace{(L \otimes I)}_{A} x + \underbrace{(L \otimes I)}_{B} \underbrace{\xi}_{\text{constant input}} \quad \xi \to 0$$

$$\dot{x} = -(L \otimes I) x \quad \Big] \text{ from before}$$

$\circ$

$\hookrightarrow \dot{x} = 0 \Rightarrow x \in \text{null}(L \otimes I)$

x has form $\mathbb{1} \otimes \tau$

$$x = \mathbb{1} \otimes \tau \qquad \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \tau \\ \vdots \\ \tau \end{bmatrix}$$

rendezvous point

$$\dot{x} = 0 \Rightarrow \underbrace{(L \otimes I)}_{A} \underbrace{\boxed{x}}_{x} = \underbrace{(L \otimes I)\xi}_{b}$$

x solves:

$x = x_0 + z$ where $z \in \text{null}(L \otimes I)$

z has the form

$x = \xi + \mathbb{1} \otimes \tau$

$\mathbb{1} \otimes \tau \to$ free parameter

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} + \begin{bmatrix} \tau \\ \tau \\ \vdots \\ \tau \end{bmatrix} \longrightarrow x_i = \xi_i + \tau$$

equation originally

arbitrary translation.