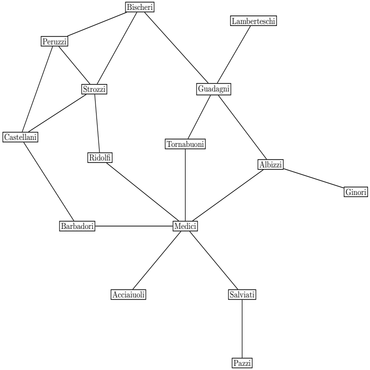
Harmonic influence in large-scale networks:

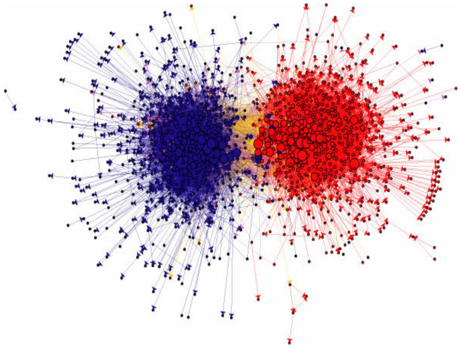
Analysis, optimization, and applications from opinion dynamics to distributed estimation

Giacomo Como[1], Fabio Fagnani[2], and Paolo Frasca[3]

MTNS 2014
Groningen, July 7, 2014

[1]Dept. of Automatic Control, Lund University
[2]Dept. of Mathematical Sciences, Politecnico di Torino
[3]Dept. of Applied Mathematics, University of Twente





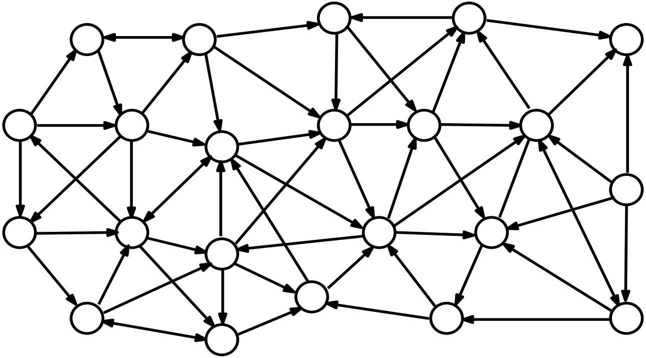
- who is more influential?
- does influence induce polarization?

* pics from Padgett and Ansell (1993) and Adamic and Glance (2005)

Plan

- centrality, opinion dynamics, stochastic matrices, mixing times
- harmonic influence, random walks, and electrical networks
- harmonic influence in large-scale networks
  - homogeneous influence
  - polarization
- optimizing harmonic influence
- distributed estimation from relative measurements

Which node is the most central?

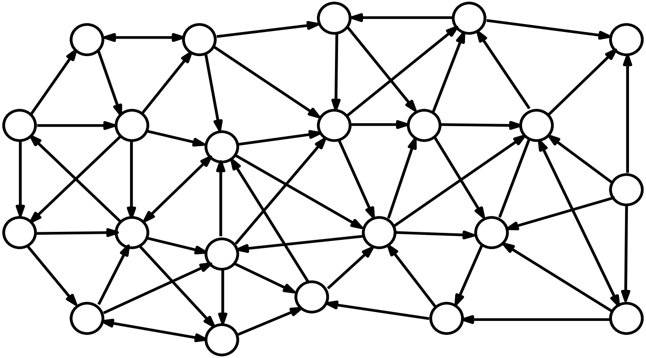


Google's Page-rank: $\sum_i \pi_i = 1$

$$\pi_i = \sum_{j \to i} \frac{1}{d_j} \pi_j$$

centrality of i out-degree of j

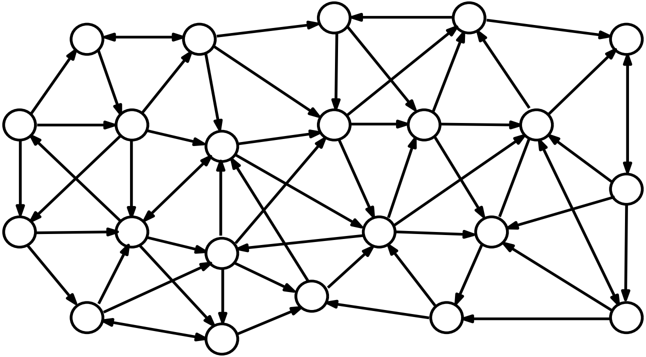
Which node is the most central?



Google's Page-rank: $\sum_i \pi_i = 1$

$$\pi_i = (1-\beta)\sum_{j\to i}\frac{1}{d_j}\pi_j + \frac{1}{n}\beta \qquad\qquad \beta \sim 0.15$$

Networks, stochastic matrices, and invariant distributions

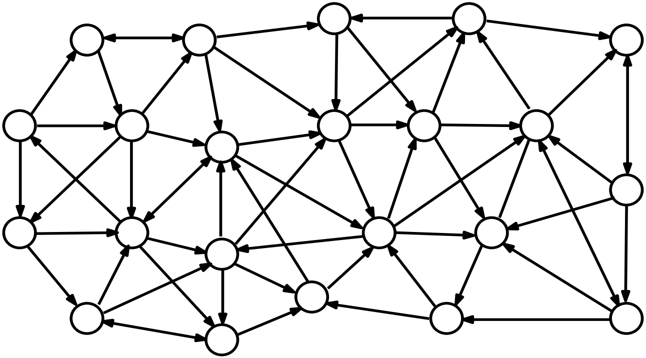


Network: $\mathcal{G} = (\mathcal{V}, \mathcal{E}, C)$, $\mathcal{V} = \{1, \ldots, n\}$

link weights $\quad C_{ij} \geq 0 \quad C_{ij} > 0 \Leftrightarrow (i, j) \in \mathcal{E}$

node degrees $d_i = \sum_j C_{ij} \qquad \overline{d} := \frac{1}{n} \sum_i d_i$

Networks, stochastic matrices, and invariant distributions

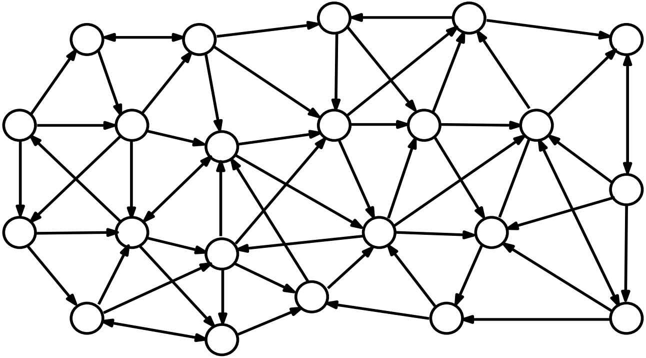


Stochastic matrix $P \in \mathbb{R}^{n \times n}$ $\qquad P_{ij} := \frac{C_{ij}}{d_i}$

stationary probability vector $\pi' = \pi' P$ $\qquad \pi' \mathbb{1} = 1$

$\mathcal{G}$ strongly connected $\implies \pi$ unique

Random walk

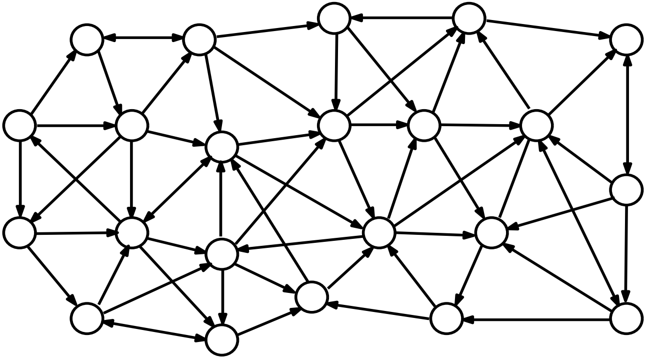


Markov chain $V(t)$ on $\mathcal{V}$ $\qquad \mathbb{P}(V(t+1)=j|V(t)=i)=P_{ij}$

$$\pi_i(t) := \mathbb{P}(V(t)=i) \qquad \pi(t+1)' = \pi(t)'P$$

$$\mathcal{G} \text{ strongly connected} \qquad \Longrightarrow \qquad \pi(t) \to \pi \quad \forall \pi(0)$$

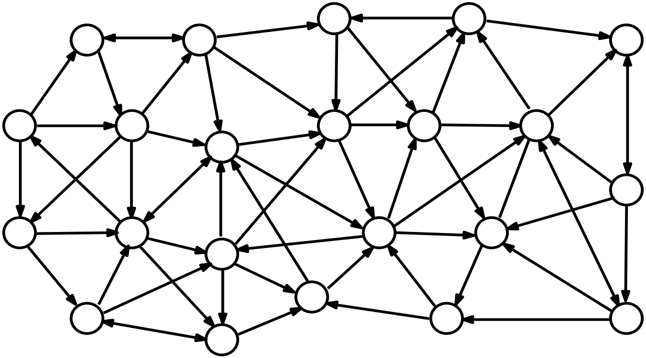
Opinion dynamics 1: distributed averaging



$$x_i(t+1) = (1-\alpha)x_i(t) + \alpha \sum_j P_{ij} x_j(t) \qquad \forall i$$

$$x(t+1) = ((1-\alpha)I + \alpha P)\, x(t) \qquad \alpha \in (0,1)$$

Opinion dynamics 1: distributed averaging

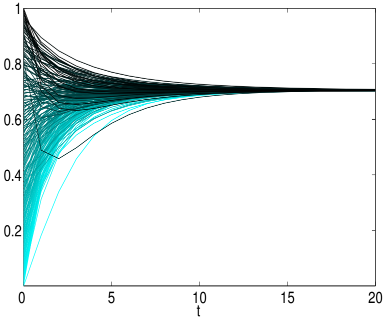


Distributed averaging as network coordination game:

$$U_i(x_i, x_{-i}) = \sum_j C_{ij}(x_i - x_j)^2$$

$$\underset{y}{\operatorname{argmin}}\, U_i(y, x_{-i}) = \sum_j P_{ij} x_j$$

Opinion dynamics 1: distributed averaging

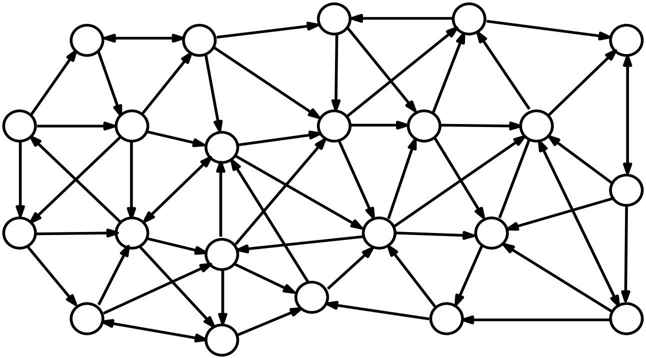


$$x(t+1) = ((1-\alpha)I + \alpha P)\, x(t)$$

$$\mathcal{G} \text{ connected} \quad \Longrightarrow \quad x(t) \to z\mathbb{1}$$

$$\text{consensus value:} \qquad z = \pi' x(0)$$

Opinion dynamics 2: Voter model

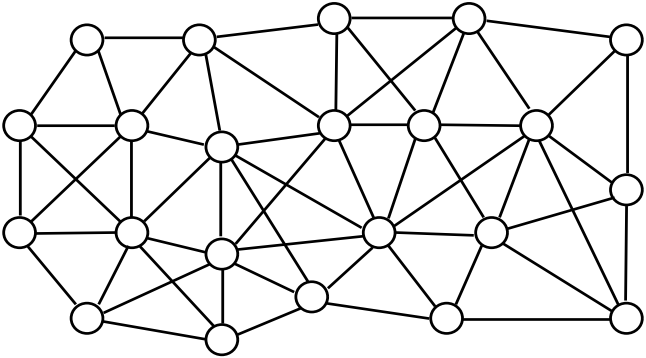


$X_i(t) \in \{0, 1\}$ i copies j at rate-P_{ij} Poisson time $\forall i$

$$\mathbb{P}\left(X(t) \xrightarrow{t\to\infty} Z\mathbb{1}\right) = 1$$

$$\mathbb{P}\left(Z = 1 | X(0)\right) = \pi' X(0)$$

Undirected networks

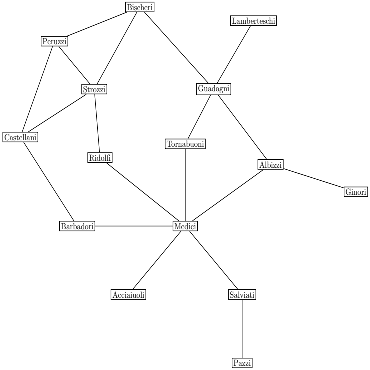


$$\mathcal{G} \text{ undirected}, \qquad C = C'$$

$$\Downarrow$$

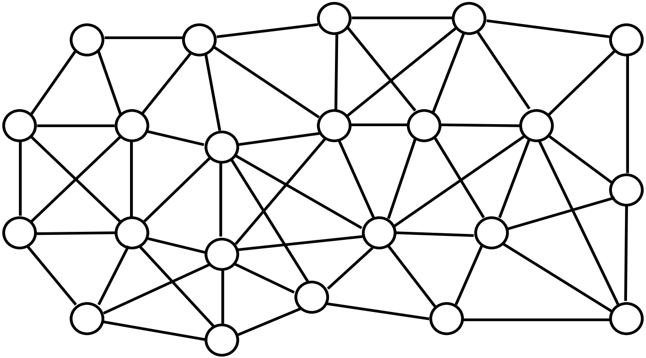
$$\pi_i = \frac{d_i}{n\bar{d}} \qquad \forall i$$

Undirected networks



$$\pi_i = \frac{d_i}{n\bar{d}} \qquad \forall i$$

Undirected networks



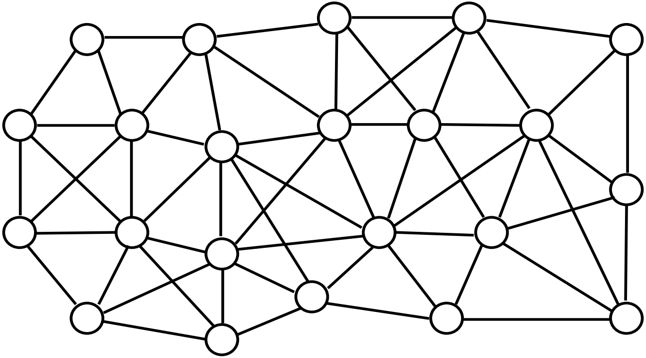
$\mathcal{G}$ undirected, $\qquad C = C'$

$$\Downarrow$$

- P reversible, i.e, $\pi_i P_{ij} = \pi_j P_{ji}$ (detailed balance)

$$V(0) \stackrel{d}{=} \pi \quad \Longrightarrow \quad (V(0), V(1), \ldots, V(t)) \stackrel{d}{=} (V(t), V(t-1), \ldots, V(0))$$

Undirected networks



$\mathcal{G}$ undirected, $\qquad C = C'$

$$\Downarrow$$

- coordination costs $U_i(x) = \sum_j C_{ij}(x_i - x_j)^2 \to$ potential game

$$\frac{\partial}{\partial x_i} U_i(x) = \frac{\partial}{\partial x_i} V(x) \qquad V(x) = \frac{1}{2} \sum_{i,j} C_{ij}(x_i - x_j)^2$$

Mixing time

$$\tau_{\text{mix}} := \inf\left\{ t \geq 0 : \max_{i,j} \frac{1}{2} ||(P^t)_{i\cdot} - (P^t)_{j\cdot}||_1 \leq \frac{1}{e} \right\}$$

- speed of convergence

$$\frac{1}{2}||\pi(t) - \pi||_1 \leq e^{-\lfloor t/\tau_{\text{mix}} \rfloor}$$

$$||x(t) - z\mathbb{1}||_\infty \leq ||x(0) - z\mathbb{1}||_\infty e^{-\lfloor \alpha t/\tau_{\text{mix}} \rfloor}$$

Mixing time

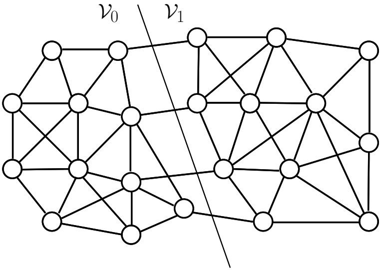
$$\tau_{\text{mix}} := \inf \left\{ t \geq 0 : \max_{i,j} \frac{1}{2} ||(P^t)_{i\cdot} - (P^t)_{j\cdot}||_1 \leq \frac{1}{e} \right\}$$

- speed of convergence

- $\mathcal{G}$ undirected $\Rightarrow$ τ_{mix} depends on conductance

$$\frac{1}{4\Phi} \leq \tau \leq \frac{8}{\Phi^2} \log \frac{n\overline{d}}{}$$

$$\Phi = \min_{0 < \sum_{i \in \mathcal{V}_0} d_i \leq \overline{d} n} \frac{\sum_{i \in \mathcal{V}_0, j \in \mathcal{V}_1} C_{ij}}{\sum_{u \in \mathcal{U}} d_u}$$



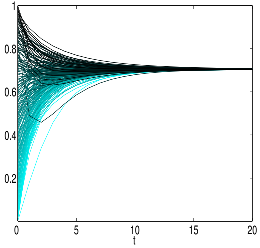
Mixing time

$$\tau_{\text{mix}} := \inf\left\{t \geq 0 : \max_{i,j} \frac{1}{2} ||(P^t)_{i\cdot} - (P^t)_{j\cdot}||_1 \leq \frac{1}{e}\right\}$$

- speed of convergence
- $\mathcal{G}$ undirected $\Rightarrow$ τ_{mix} depends on conductance
- small τ_{mix} $\Rightarrow$ robustness of π to perturbations of P

  More on this in the talk on Tue, July 8, at 10:30, in A.7

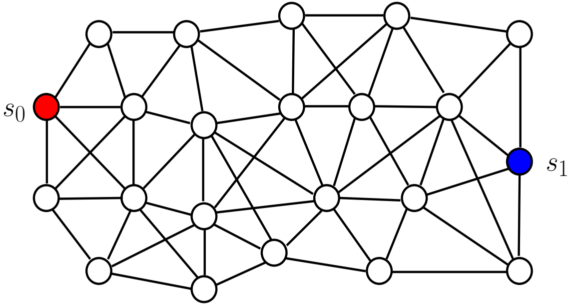
Consensus vs disagreement

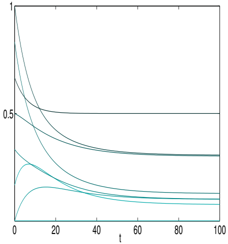


"Since universal ultimate agreement is an ubiquitous outcome of a very broad class of mathematical models, we are naturally led to inquire what on earth one must assume in order to generate ..." (Abelson '64)

"If people tend to become more alike in their beliefs, attitudes, and behavior as they interact, why do not such differences eventually disappear?" (Axelrod '97)

Opinion dynamics with stubborn agents



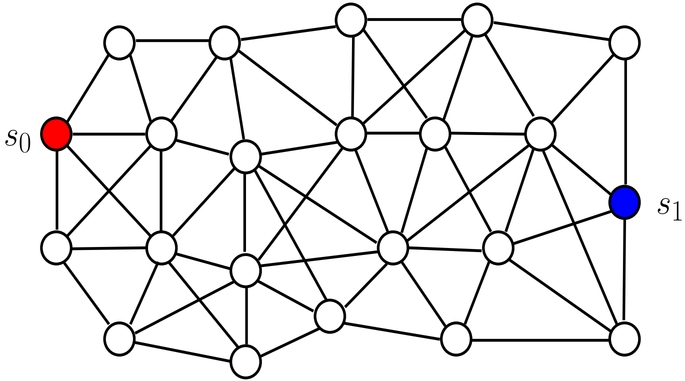


$$x_{s_0}(t) \equiv 0 \qquad x_{s_1}(t) \equiv 1$$

$$x_i(t+1) = (1-\alpha)x_i(t) + \alpha \sum_j P_{ij} x_j(t) \qquad i \neq s_0, s_1$$

$$x(t) \overset{t \to \infty}{\longrightarrow} x$$

Stubborn nodes and Harmonic influence

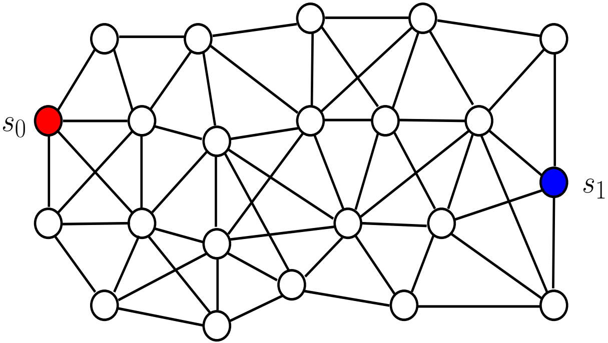


Harmonic influence vector: $x \in \mathbb{R}^n$ unique solution of

$$x_i = \sum_j P_{ij} x_j \qquad i \neq s_0, s_1$$

$$x_{s_0} = 0 \qquad x_{s_1} = 1$$

Stubborn nodes and Harmonic influence



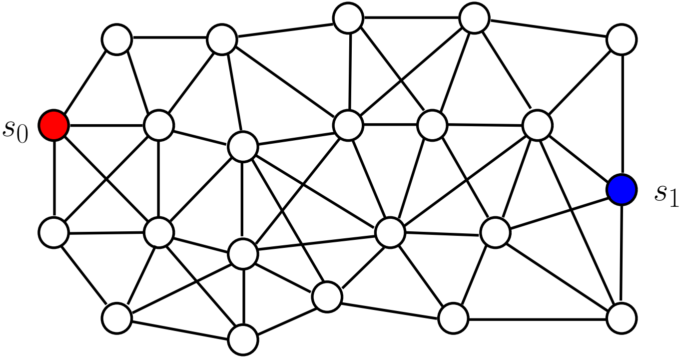
Harmonic influence vector as Nash equilibrium

$$x_i \in \underset{y}{\operatorname{argmin}}\, U_i(y, x_{-i})$$

$$U_i(x_i, x_{-i}) = \sum_j C_{ij}(y - x_j)^2 \qquad i \neq s_0, s_1$$

$$U_{s_0}(x) = x_{s_0}^2 \qquad U_{s_1}(x) = (x_{s_1} - 1)^2$$

Opinion dynamics with stubborn nodes 2: voter model



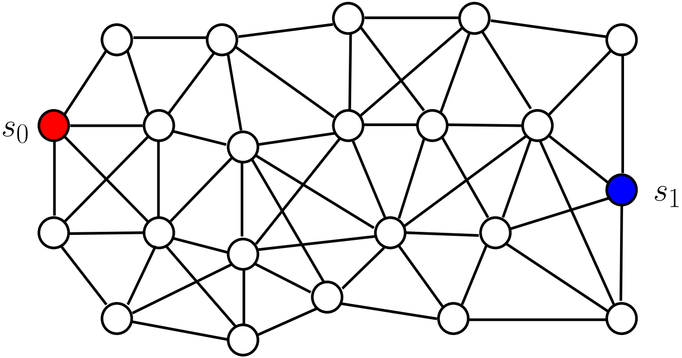
$X_{s_0}(t) \equiv 0 \qquad X_{s_1}(t) \equiv 1 \qquad X_i(t) \in \{0, 1\} \quad i \neq s_0, s_1$

$i \neq s_0, s_1$ copies j at rate-C_{ij} Poisson time

$X(t) \xrightarrow{d} X \qquad \mathbb{P}(X_i = 1) = x_i \qquad$ (note: fluctuations persist)

Random walk interpretation

$$\mathbb{P}(V(t+1) = j | V(t) = i) = P_{ij}$$

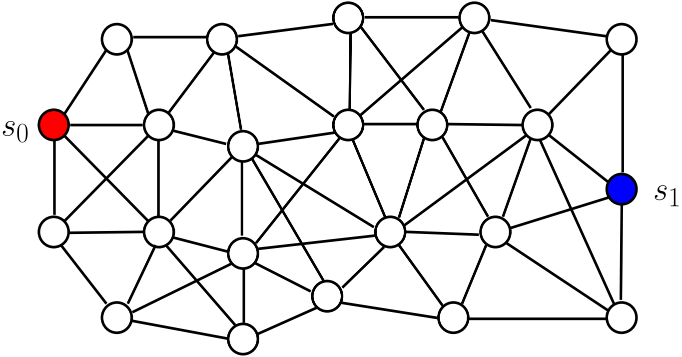


$T_j :=$ hitting time on j

$$x_i = \mathbb{P}(T_{s_1} < T_{s_0} | V(0) = i)$$

Random walk interpretation

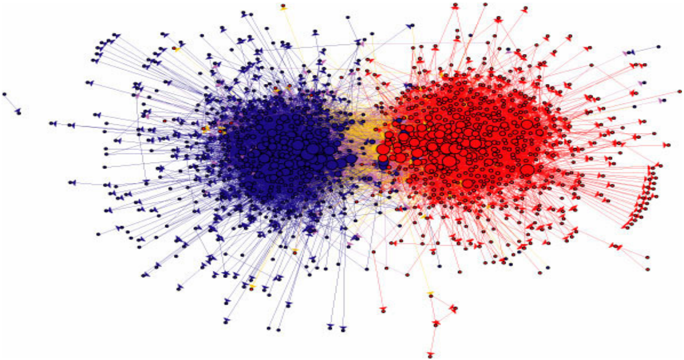
$$\mathbb{P}(V(t+1) = j | V(t) = i) = P_{ij}$$



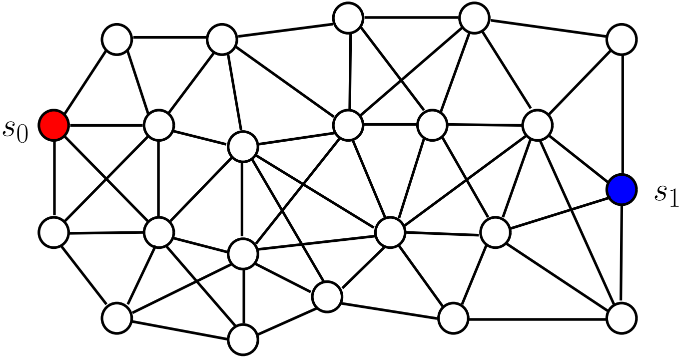
$T_j :=$ hitting time on j $\qquad \tau_i^j := \mathbb{E}[T_j | V(0) = i]$

$$x_i = \overline{x} + \frac{\tau_{s_0}^i - \tau_{s_1}^i}{\tau_{s_1}^{s_0} + \tau_{s_0}^{s_1}} \qquad \overline{x} = \frac{\sum_i d_i x_i}{\sum_j d_j} = \frac{\tau_{s_0}^{s_1}}{\tau_{s_1}^{s_0} + \tau_{s_0}^{s_1}}$$

How many stubborn nodes can polarize a network?



How many stubborn nodes can polarize a network?



fluidity $\qquad \Phi := \dfrac{n d_* / \overline{d}}{\tau_{\mathrm{mix}}(d_{s_0} + d_{s_1})}$

$\tau_{\mathrm{mix}} :=$ mixing time of P $\qquad d_* := \min_i d_i \qquad \overline{d} = \frac{1}{n}\sum_i d_i$

How many stubborn nodes can polarize a network?

$$\text{fluidity} \qquad \Phi := \frac{n d_* / \overline{d}}{\tau_{\text{mix}}(d_{s_0} + d_{s_1})}$$

Theorem [Acemoglu, Como, Fagnani, Ozdaglar, MOR 2013]

$$\frac{1}{n} \left| \{ i : |x_i - \overline{x}| \geq \varepsilon \} \right| \leq \frac{1}{\Phi \varepsilon} \qquad \forall \varepsilon > 0$$

Corollary:

Highly fluid: $\Phi \overset{n \to \infty}{\longrightarrow} \infty$ $\Longrightarrow$ homogeneous influence: x almost constant on $\mathcal{V}$

How many stubborn nodes can polarize a network?

$$\Phi := \frac{n d_* / \overline{d}}{\tau_{\text{mix}}(d_{s_0} + d_{s_1})} \qquad \overline{x} = \frac{1}{n} \sum_i x_i$$

- if $\tau_{\text{mix}} \le K \log^k n$ and $\overline{d} \le K' d_*$ then

$$(d_{s_0} + d_{s_1}) \le K'' n^{1-\varepsilon} \implies \Phi \to \infty$$

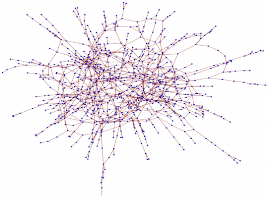
Highly fluid networks

(s_0, s_1 obtained by merging nodes from $\mathcal{S} \subseteq \mathcal{V}$)

- Connected Erdös-Rényi

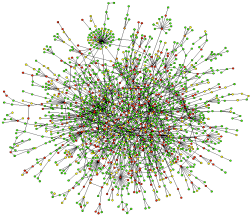
  $\mathcal{G} = ER(n, p = \frac{c}{n} \log n)\,, \quad c > 1$

  $$|\mathcal{S}| = o\left(\frac{n}{\log n}\right) \implies \text{highly fluid w.h.p.}$$



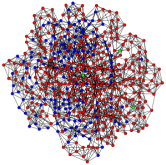
- Preferential attachment [Barabasi'99]

  $$\sum_s d_s = o\left(\frac{n}{\log n}\right) \implies \text{highly fluid w.h.p}$$



- Small world [Watts&Strogatz'98]

  $$\sum_s d_s = o\left(\frac{n}{\log^3 n}\right) \implies \text{highly fluid w.h.p.}$$



Homogeneous influence vs uncorrelated opinions

In voter model with stubborn agents $X(t) \xrightarrow{d} X$.

Persistent fluctuations.

ergodic aggregate belief	$\overline{X} := \frac{1}{n}\sum_i X_i$
mean square disagreement	$\Delta^2 := \frac{1}{2n^2}\sum_{i,j}\mathbb{E}\left[(X_i - X_j)^2\right]$

Proposition: highly fluid $\Longrightarrow$ $\Delta^2 + \mathrm{Var}[\overline{X}] = \sigma^2 + o(1)$

Homogeneous influence vs uncorrelated opinions (cont'd)

Theorem: in highly fluid networks

$$\frac{\tau}{\pi(\mathcal{S})}\frac{\overline{d^2}}{n\overline{d}^2} \xrightarrow{n\to\infty} 0 \quad \Longrightarrow \quad \mathrm{Var}[\overline{X}] \to 0, \quad \Delta^2 = \sigma^2 + o(1)$$

Homogeneous influence vs uncorrelated opinions (cont'd)

Theorem: in highly fluid networks

$$\frac{\tau}{\pi(\mathcal{S})}\frac{\overline{d^2}}{n\bar{d}^2} \xrightarrow{n\to\infty} 0 \quad \Longrightarrow \quad \mathrm{Var}[\overline{X}] \to 0, \quad \Delta^2 = \sigma^2 + o(1)$$

- Connected Erdös-Rényi

  $$\omega(\log n) = |\mathcal{S}| = o\Big(\frac{n}{\log n}\Big)$$

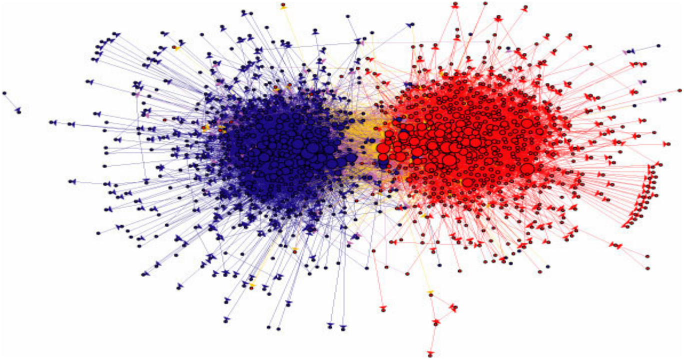
- Preferential attachment

  $$\omega(\log n) = \sum_s d_s = o\Big(\frac{n}{\log n}\Big)$$

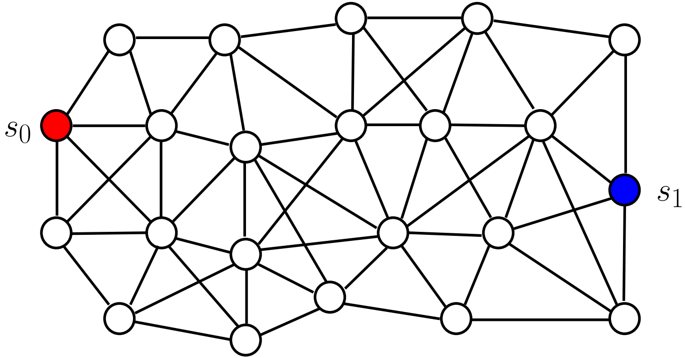
- Small world

  $$\omega(\log^3 n) = \sum_s d_s = o\Big(\frac{n}{\log^3 n}\Big)$$

How many stubborn nodes can polarize a network?



Harmonic influence in undirected networks

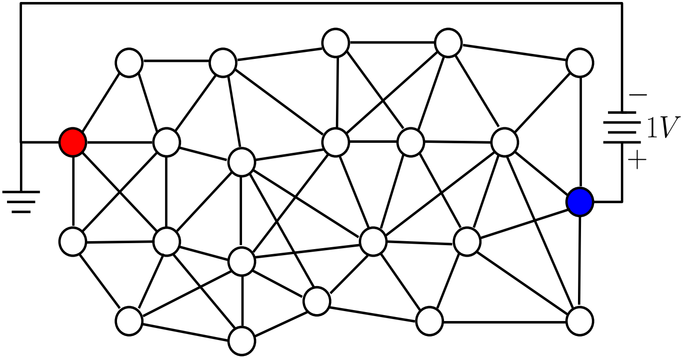


- $\mathcal{G}$ undirected, $C = C'$

$$x_i = \sum\nolimits_j P_{ij} x_j \quad i \neq s_0, s_1 \qquad \Longleftrightarrow \qquad x = \underset{\substack{y \in \mathbb{R}^n: \\ y_{s_0}=0 \;\; y_{s_1}=1}}{\operatorname{argmin}} \frac{1}{2} \sum_{i,j} C_{ij} (y_i - y_j)^2$$

$$x_{s_0} = 0 \qquad x_{s_1} = 1$$

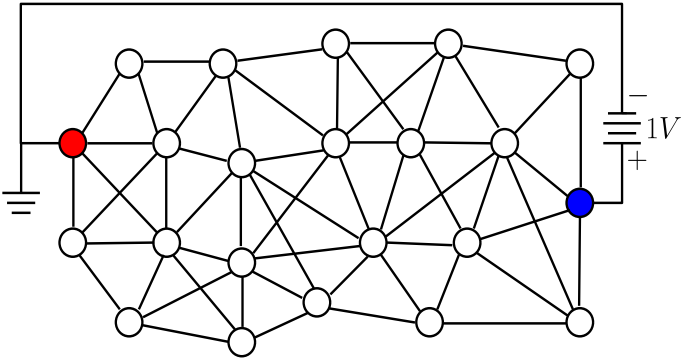
Electrical network interpretation



$$x = \underset{\substack{y\in\mathbb{R}^n:\\ y_{s_0}=0 \;\; y_{s_1}=1}}{\operatorname{argmin}} \frac{1}{2}\sum_{i,j} C_{ij}(y_i - y_j)^2$$

- C_{ij} = conductance of link $\{i,j\}$
- x_i = voltage at node i

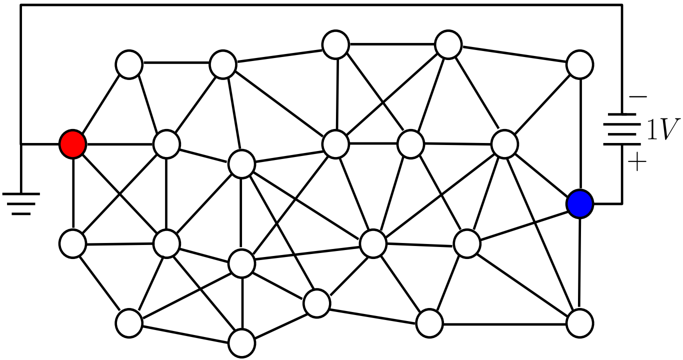
Electrical network interpretation



$$x = \operatorname*{argmin}_{\substack{y \in \mathbb{R}^n: \\ y_{s_0}=0 \;\; y_{s_1}=1}} \frac{1}{2} \sum_{i,j} C_{ij} (y_i - y_j)^2$$

- Ohm: $\iota(i,j) := C_{ij}(x_i - x_j)$ current flow i to j
- $\Rightarrow$ Kirchoff: $\sum_j \iota(i,j) = 0 \;\forall i \neq s_0, s_1$

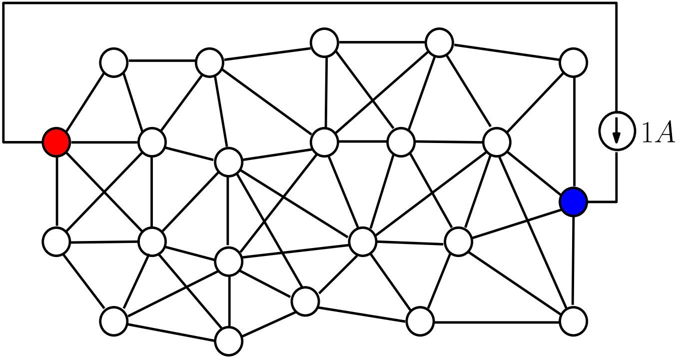
Effective resistance



$$R(s_0 \leftrightarrow s_1) := \min_{\substack{y \in \mathbb{R}^n: \\ y_{s_0}=0 \;\; y_{s_1}=1}} \frac{1}{2} \sum_{i,j} C_{ij}(y_i - y_j)^2$$

- $C_{ij}(x_i - x_j)^2 = \iota(i,j)(x_i - x_j) = \dfrac{\iota(i,j)^2}{C_{i,j}}$ heat dissipation on $\{i,j\}$

Thompson's variational principle



$$R(s_0 \leftrightarrow s_1) := \min_{\substack{\theta \text{ unitary flow} \\ \text{from } s_1 \text{ to } s_0}} \frac{1}{2} \sum_{i,j} \frac{1}{C_{ij}} \theta^2(i,j)$$

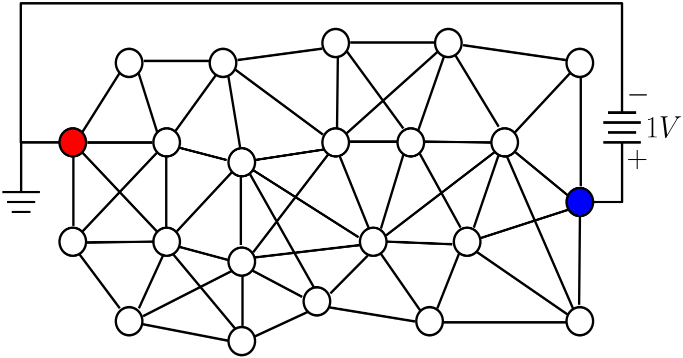
achieved by $\theta(i,j) = \dfrac{\iota(i,j)}{\sum_j \iota(s_1, j)}$

Computation/estimation of effective resistances

- series and parallel laws
- glueing nodes does not increase effective resistance
- removing links does not decrease effective resistance
- Laplacian $L = \mathrm{diag}(C\mathbb{1}) - C$ Green function: $Z = \mathrm{pinv}(L)$

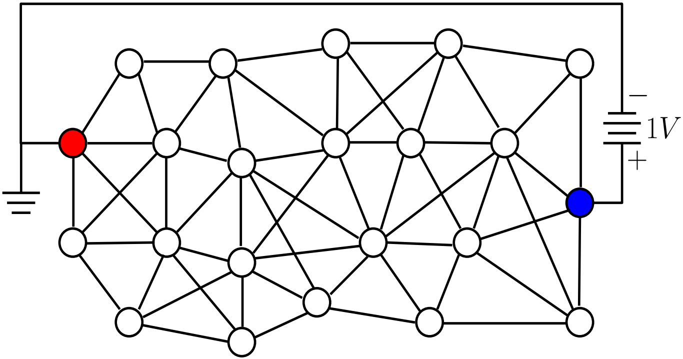
$$R(i \leftrightarrow j) = Z_{ii} + Z_{jj} - 2Z_{ij}$$

Effective resistance and local influence



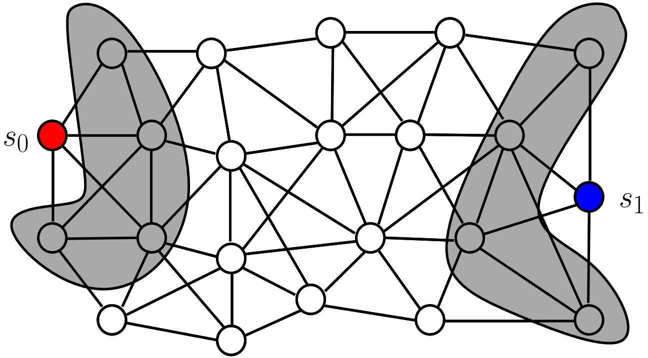
$$\underbrace{\sum_j \iota(j, s_0 j)}_{\text{current in } s_0} = \frac{x_{s_1} - x_{s_0}}{R(s_0 \leftrightarrow s_1)} = \underbrace{\sum_j \iota(s_1, j)}_{\text{current to } s_1}$$

Effective resistance and local influence



$$\underbrace{\sum_j C_{s_0 j}(x_j - x_{s_0})}_{\text{current in } s_0} \;=\; \frac{x_{s_1} - x_{s_0}}{R(s_0 \leftrightarrow s_1)} \;=\; \underbrace{\sum_j C_{s_1 j}(x_{s_1} - x_j)}_{\text{current to } s_1}$$

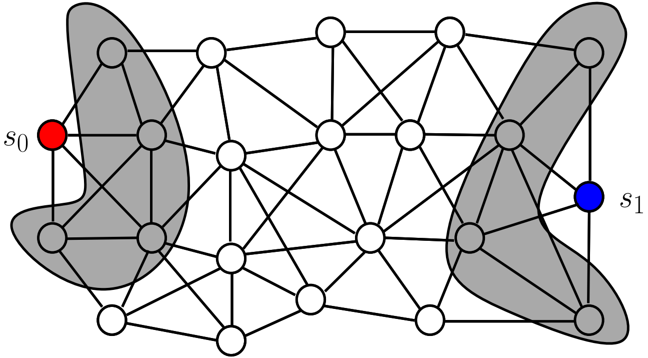
Effective resistance and local influence



$y_0 := \sum_j P_{s_0 j} x_j$ $\qquad$ $y_1 := \sum_j P_{s_1 j} x_j$

$$\underbrace{d_{s_0} y_0}_{\text{current in } s_0} = \frac{1}{R(s_0 \leftrightarrow s_1)} = \underbrace{d_{s_1}(1 - y_1)}_{\text{current to } s_1}$$

Effective resistance and local influence

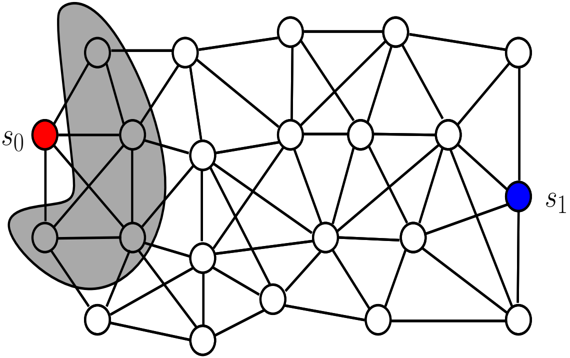


$$y_0 := \sum_j P_{s_0 j} x_j \qquad\qquad y_1 := \sum_j P_{s_1 j} x_j$$

$$d_{s_0} y_0 \quad = \quad d_{s_1}(1 - y_1)$$

- $\dfrac{d_{s_0}}{d_{s_1}} \to \infty \quad \Longrightarrow \quad y_0 \to 0$
- $\dfrac{d_{s_1}}{d_{s_0}} \to \infty \quad \Longrightarrow \quad y_1 \to 1$

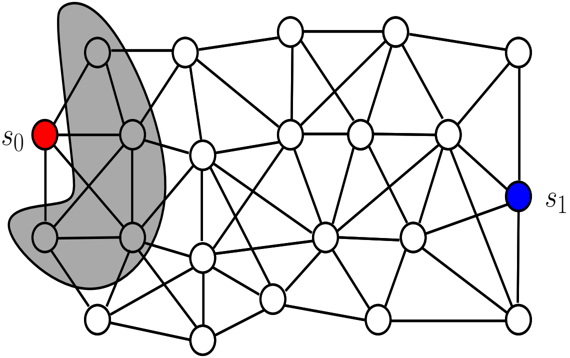
From local to global influence



Recall: highly fluid $\implies \frac{1}{n}|\{i : |x_i - \overline{x}| \geq \varepsilon\}| \to 0$

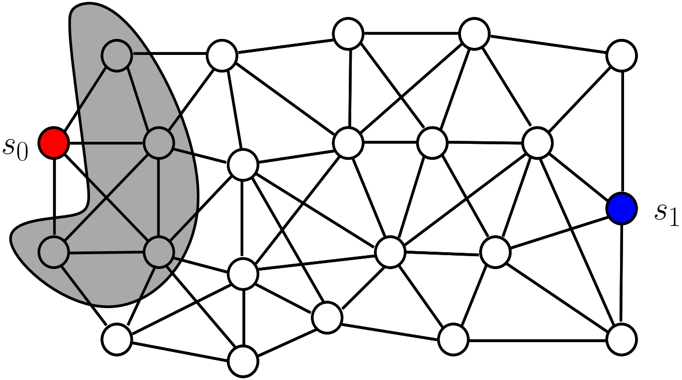
- $\frac{d_{s_0}}{d_{s_1}} \to \infty$ AND highly fluid $\overset{?}{\implies}$ $\frac{1}{n}|\{i : x_i \geq \varepsilon\}| \to 0$

From local to global influence



- $\dfrac{d_{s_0}}{d_{s_1}} \to \infty$ AND highly fluid $\not\Rightarrow$ $\dfrac{1}{n}|\{i : x_i \geq \varepsilon\}| \to 0$

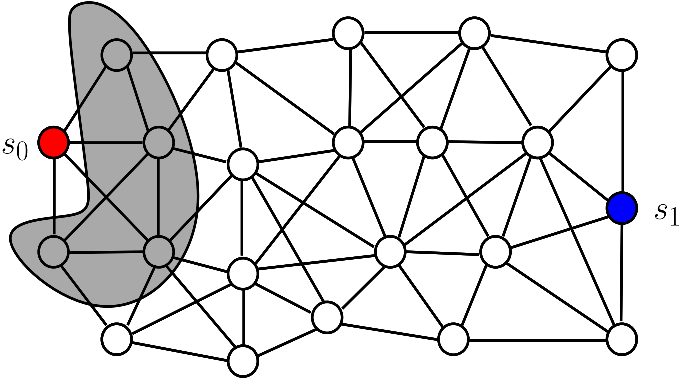
Escape probability



T_i^+ = return time in i for random walk

$$\zeta_i := \sup_{k\geq 0} \frac{\mathbb{P}(T_i^+ > k\tau_{\text{mix}} | V(0) = i) - 2e^{-k}}{1 + k\tau_{\text{mix}}\pi_i} = \text{“}\mathbb{P}(T_i^+ \gg \tau_{\text{mix}})\text{''}$$

Escape probability



$\liminf \zeta_i > 0$ if there is positive drift away from i

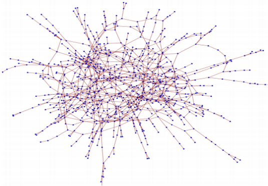
Escape probability in large-scale networks

(i obtained by merging nodes from random $\mathcal{S} \subseteq \mathcal{V}$)

- Connected Erdös-Rényi

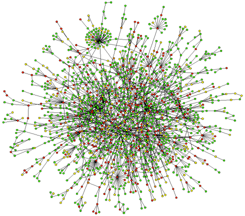
  $\mathcal{G} = ER(n, p = \frac{c}{n} \log n)\,, \quad c > 1$

$$|\mathcal{S}| = O(n^{1-\varepsilon}) \implies \liminf \zeta_s > 0 \text{ w.h.p.}$$



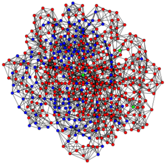
- Preferential attachment [Barabasi'99]

$$\sum_s d_s = O(n^{1-\varepsilon}) \implies \liminf \zeta_s > 0 \text{ w.h.p.}$$

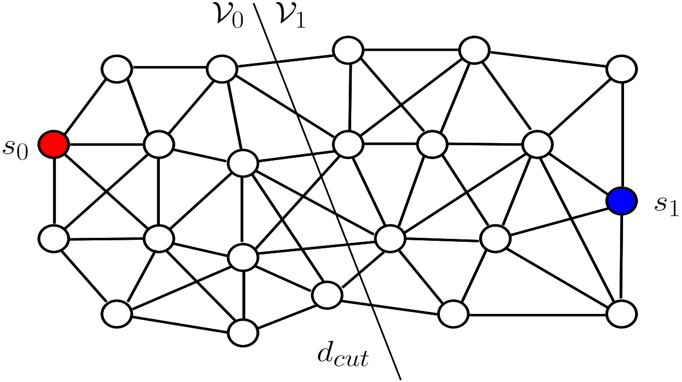


- Small world [Watts&Strogatz'98]

$$\sum_s d_s = O(n^{1-\varepsilon}) \implies \liminf \zeta_i > 0 \text{ w.h.p.}$$

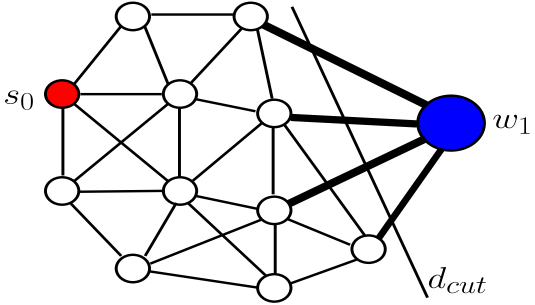


Sufficient condition for polarization



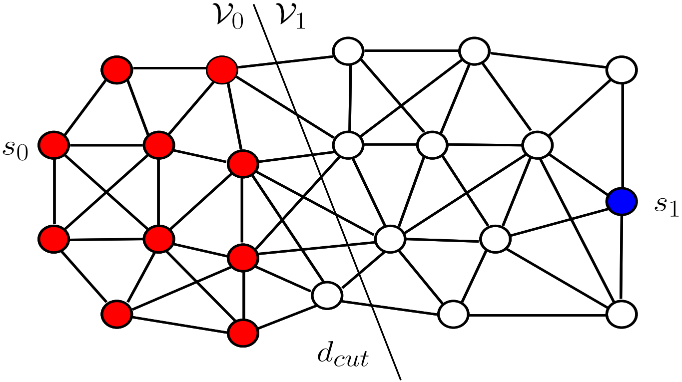
Relative cut $\mathcal{V} = \mathcal{V}_0 \cup \mathcal{V}_1$.

Sufficient condition for polarization



Relative cut $\mathcal{V} = \mathcal{V}_0 \cup \mathcal{V}_1$. $\mathcal{G}_0 := \mathcal{G}$ with $\mathcal{V}_1$ collapsed in one node

Sufficient condition for polarization [ACFO,'14]

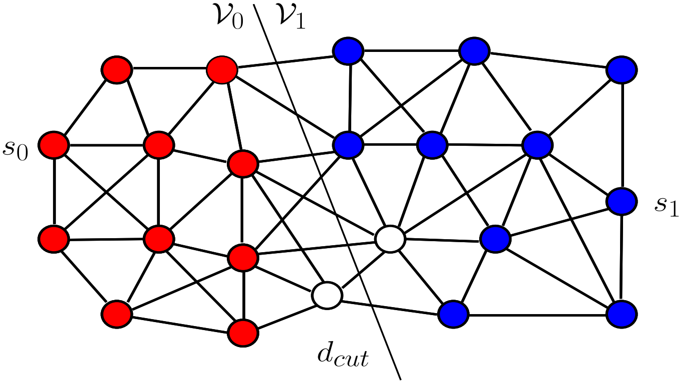


Relative cut $\mathcal{V} = \mathcal{V}_0 \cup \mathcal{V}_1$. $\quad \mathcal{G}_0 := \mathcal{G}$ with $\mathcal{V}_1$ collapsed in one node

Theorem 1: $\quad d_{cut}/d_{s_0} \to 0, \qquad \mathcal{G}_0$ highly fluid, $\qquad \liminf \zeta^0_{s_0} > 0$

$$\Longrightarrow \qquad \frac{1}{|\mathcal{V}_0|} \left|\{i \in \mathcal{V}_0 : x_i > \varepsilon\}\right| \to 0 \qquad \forall \varepsilon > 0$$

Sufficient condition for polarization

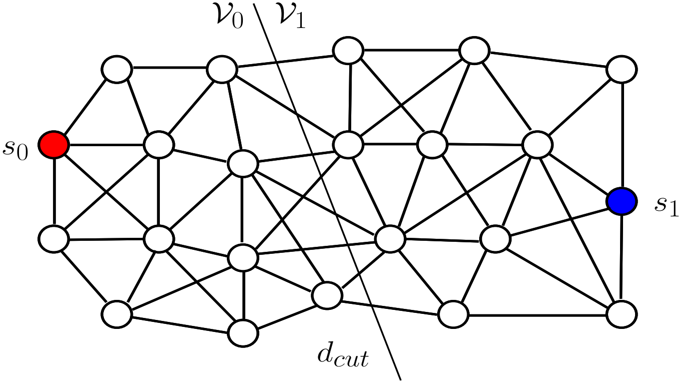


Relative cut $\mathcal{V} = \mathcal{V}_0 \cup \mathcal{V}_1$. $\mathcal{G}_1 := \mathcal{G}$ with $\mathcal{V}_0$ collapsed in one node

Theorem 1: $d_{cut}/d_{s_1} \to 0$, $\mathcal{G}_1$ highly fluid, $\liminf \zeta^1_{s_1} > 0$

$$\implies \quad \frac{1}{|\mathcal{V}_1|} \left|\{i \in \mathcal{V}_1 : x_i < 1 - \varepsilon\}\right| \to 0 \qquad \forall \varepsilon > 0$$

Sufficient condition for (weakly) homogeneous influence

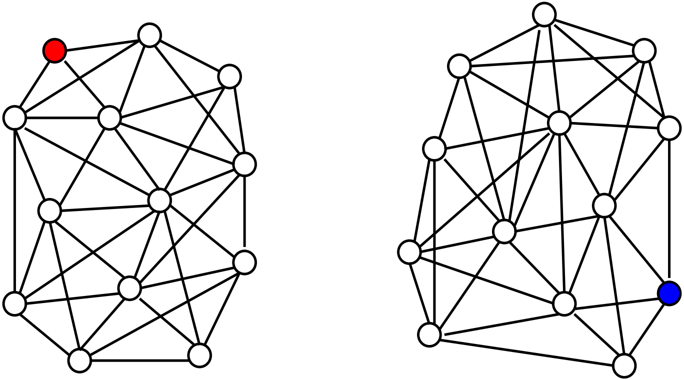


Relative cut $\mathcal{V} = \mathcal{V}_0 \cup \mathcal{V}_1$. $\quad z_0 := \max_{i \in \mathcal{V}_0} x_i \quad z_1 := \min_{i \in \mathcal{V}_1} x_i$

Theorem 2: $\dfrac{d_{s_1} + d_{s_1}}{d_{cut}} \to 0$, $\mathcal{G}_0, \mathcal{G}_1$ highly fluid, $\liminf \zeta^0_{w_1} \zeta^1_{w_0} > 0$

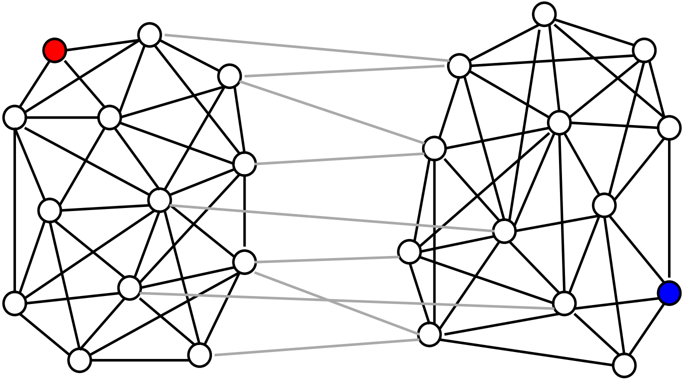
$$\Longrightarrow \qquad \frac{1}{n} \left| \{ i \in \mathcal{V} : \, z_1 - \varepsilon < x_i < z_0 + \varepsilon \} \right| \to 1 \qquad \forall \varepsilon > 0$$

‘Phase-transition’



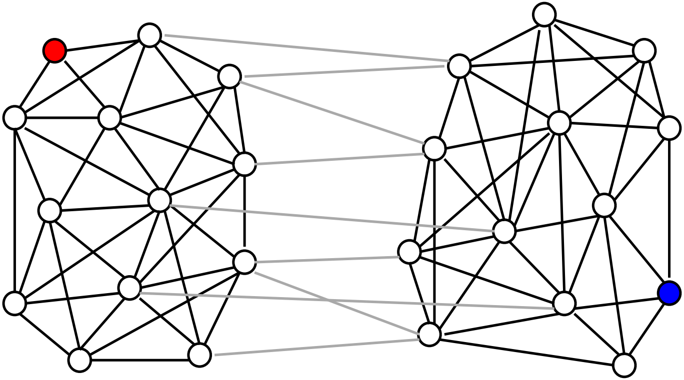
- highly fluid $\mathcal{G}_0 = (\mathcal{V}_0, \mathcal{E}_0)$, $\mathcal{G}_1 = (\mathcal{V}_0, \mathcal{E}_0)$ with $s_0 \in \mathcal{V}_0$, $s_1 \in \mathcal{V}_1$

‘Phase-transition’



- highly fluid $\mathcal{G}_0 = (\mathcal{V}_0, \mathcal{E}_0)$, $\mathcal{G}_1 = (\mathcal{V}_0, \mathcal{E}_0)$ with $s_0 \in \mathcal{V}_0$, $s_1 \in \mathcal{V}_1$
- connect every $i \in \mathcal{V}_0$ with $j \in \mathcal{V}_1$ indep. with probability α

‘Phase-transition’



Theorem

- $$\alpha << \frac{d_{s_0} + d_{s_1}}{n^2} \implies \text{polarization}$$
- $$\alpha >> \frac{d_{s_0} + d_{s_1}}{n^2} \implies \text{(weakly) homogeneous influence}$$

Plan

- centrality, opinion dynamics, stochastic matrices, mixing times
- harmonic influence, random walks, and electrical networks
- harmonic influence in large-scale networks
  - homogeneous influence
  - polarization
- optimizing harmonic influence
- distributed estimation from relative measurements

References

- Aldous and Fill, Reversible Markov chains and random walks on graphs, monograph in preparation, 2002
- Levin, Peres, and Wilmer, Markov chains and mixing times, American Mathematical Society, 2010
- Doyle and Snell, Random walks and electric networks, The Mathematical Association of America, 1984
- Acemoglu, Como, Fagnani, and Ozdaglar, "Opinion fluctuations and disagreement in social networks", Mathematics of Operation Research, 38 (1), pp. 1–27, 2013
- Acemoglu, Como, Fagnani, and Ozdaglar, "Harmonic influence in large-scale networks", preprint, presented at NetEcon, 2014

Harmonic influence in large-scale networks:

Analysis, optimization, and applications from opinion dynamics to distributed estimation

Giacomo Como[1], **Fabio Fagnani**[2], and Paolo Frasca[3]

MTNS 2014
Groningen, July 7, 2014

[1]Dept. of Automatic Control, Lund University
[2]Dept. of Mathematical Sciences, Politecnico di Torino
[3]Dept. of Applied Mathematics, University of Twente

Part II: Optimal placement problems.

- What are the most influential nodes in a network?
  - What are the nodes to conquer to maximize influence?
  - What are the nodes to defend to minimize the effect of possible future invasions?
- In this talk, we will focus on these topics using the Harmonic centrality as a measure of influence.

Summary

- Harmonic influence
- Two optimality problems
- Structural properties.
- A message-passing recursive algorithm for trees
- Theoretical results for general graphs
- Simulations

Harmonic influence

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ graph. $n = |\mathcal{V}|$. $\mathcal{S} = \mathcal{S}_0 \cup \mathcal{S}_1 \subseteq \mathcal{V}$ stubborn agents.

P stochastic matrix on $\mathcal{G}$ (e.g. SRW)

Asymptotic opinions under a consensus dynamics are characterized by the Laplace equation with boundary conditions:

$$x_i = \sum_{j \in \mathcal{V}} P_{ij} x_j \,, \quad \forall i \in \mathcal{V} \setminus \mathcal{S} \,, \quad x_s = a \;\; \forall s \in S_a$$

Harmonic influence centrality (HIC):

$$H(\mathcal{S}_0, \mathcal{S}_1) := \bar{x} = \frac{1}{n} \sum_{i \in \mathcal{V}} x_i$$

Two optimality problems

$H(\mathcal{S}_0, \mathcal{S}_1)$ Harmonic influence centrality.

Two problems:

1. Given $\mathcal{S}_0 \subseteq \mathcal{V}$, find the best placement for $\mathcal{S}_1$

$$\max_{\mathcal{S}_1 \subseteq \mathcal{V} \setminus \mathcal{S}_0} H(\mathcal{S}_0, \mathcal{S}_1)$$

   ($|\mathcal{S}_1| = n_1$ assigned)

2. Find the best placement for $\mathcal{S}_0$ assuming he can choose first

$$\min_{\mathcal{S}_0 \subseteq \mathcal{V}} \max_{\mathcal{S}_1 \subseteq \mathcal{V} \setminus \mathcal{S}_0} H(\mathcal{S}_0, \mathcal{S}_1)$$

   ($|\mathcal{S}_i| = n_i$ assigned)

Two optimality problems

Standing assumptions in this part:

- $\mathcal{G}$ undirected, P reversible (e.g. SRW).
- Electrical interpretation:
  - x_i voltage at node i.
  - $R(i \leftrightarrow j)$ effective resistance.
- A useful formula:

$$x_j = \frac{1}{2} + \frac{R(\mathcal{S}_0 \leftrightarrow j) - R(\mathcal{S}_1 \leftrightarrow j)}{2R(\mathcal{S}_0 \leftrightarrow \mathcal{S}_1)}$$

  Proof follows from classical electrical networks tools (Green function ...)

$$H(\mathcal{S}_0, \mathcal{S}_1) = \frac{1}{2} + \frac{\overline{R}(\mathcal{S}_0) - \overline{R}(\mathcal{S}_1)}{2R(\mathcal{S}_0 \leftrightarrow \mathcal{S}_1)}$$

where $\overline{R}(\mathcal{S}_a) := \frac{1}{|\mathcal{V}|} \sum_{j \in \mathcal{V}} R(\mathcal{S}_a \leftrightarrow j)$.

Some preliminary results

A special case: $\mathcal{S}_0 = \{s_0\}$; $\mathcal{S}_1 = \{s_1\}$ both singletons.

$$H(s_0, s_1) = \frac{1}{2} + \frac{\overline{R}(s_0) - \overline{R}(s_1)}{2R(s_0 \leftrightarrow s_1)}$$

The optimal placement for s_0:

- $\operatorname*{argmin}_{s_0 \in \mathcal{V}} \max_{s_1 \in \mathcal{V} \setminus \{s_0\}} H(s_0, s_1) = \operatorname*{argmin}_{s_0 \in \mathcal{V}} \overline{R}(s_0)$
- If s_0 is optimal, $H(s_0, s_1) \leq 1/2$ for every choice of s_1
- If s_0 is optimal and $\exists s_1 \neq s_0$ s.t. $\overline{R}(s_0) = \overline{R}(s_1)$, then $H(s_0, s_1) = 1/2$

To choose first is better!

Some preliminary results

A special case: $\mathcal{S}_0 = \{s_0\}$; $\mathcal{S}_1 = \{s_1\}$ both singletons.

$$H(s_0, s_1) = \frac{1}{2} + \frac{\overline{R}(s_0) - \overline{R}(s_1)}{2R(s_0 \leftrightarrow s_1)}$$

The optimal placement for s_1:

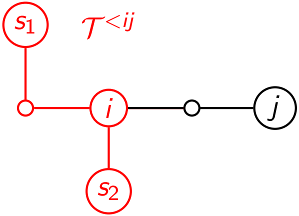
- If s_0 is optimally placed
  - and there is not another minimum for $\overline{R}$, the best placement for s_1 is a trade-off between minimizing $\overline{R}(s_1)$ with $s_1 \neq s_0$ and maximizing $R(s_0 \leftrightarrow s_1)$. In this case $H < 1/2$
- If s_0 is not optimally placed,
  - the optimal placement for s_1 is always such that $\overline{R}(s_1) < \overline{R}(s_0)$ so that $H > 1/2$
  - the optimal placement for s_1 is a trade-off between minimizing $\overline{R}(s_1)$ with $s_1 \neq s_0$ and minimizing $R(s_0 \leftrightarrow s_1)$.

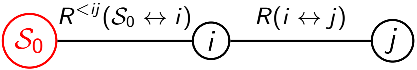
Two possible different strategies: 'close to' or 'far from' s_1.

The optimal placement for s_1 is a more difficult problem!

When the network $\mathcal{T} = (\mathcal{V}, \mathcal{E})$ is a tree

- $i, j \in \mathcal{V}$, $\mathcal{T}^{<ij}$ subtree from node i including all branches not containing j.
- $R^{<ij}(\mathcal{S}_0 \leftrightarrow i)$ eff. resistance between $\mathcal{S}_0$ and i in $\mathcal{T}^{<ij}$.





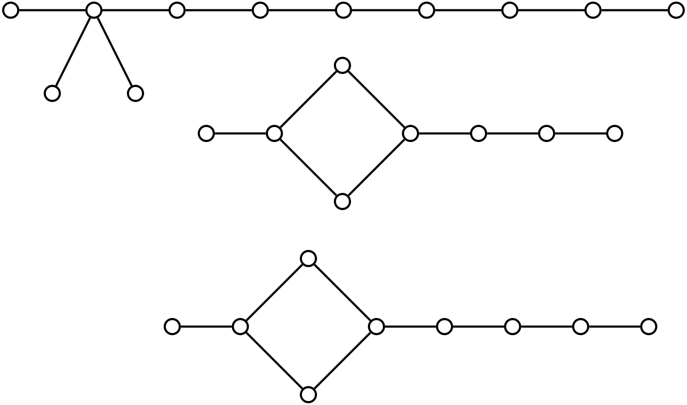
- A natural iterative computational structure: $i \in \mathcal{V}$, $j \in N_i$.

$$R^{<ij}(\mathcal{S}_0 \leftrightarrow i) = \frac{1}{\sum_{k \in N_i \setminus \{j\}} R^{<kj}(\mathcal{S}_0 \leftrightarrow k) + 1}$$

$$R(\mathcal{S}_0 \leftrightarrow i) = \frac{1}{\sum_{k \in N_i} R^{<kj}(\mathcal{S}_0 \leftrightarrow k) + 1}$$

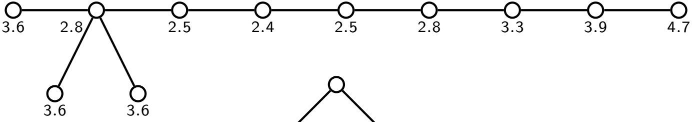
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

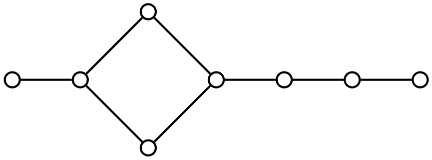
The role of the mean effective resistances $\overline{R}$.

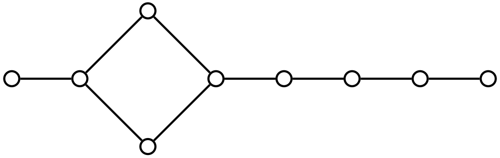


Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

The role of the mean effective resistances $\overline{R}$.

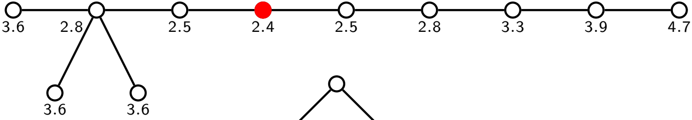


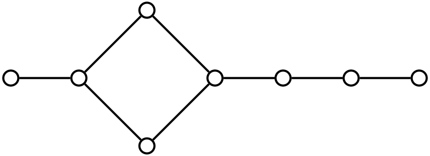


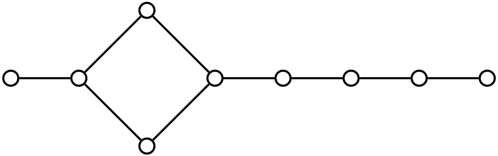


Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

The role of the mean effective resistances $\overline{R}$.

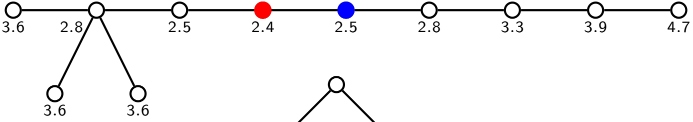


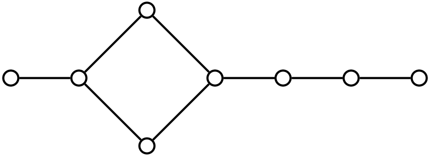


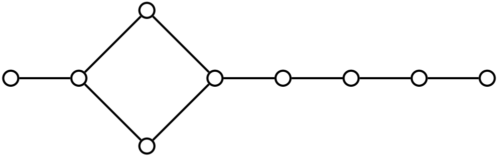


Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

The role of the mean effective resistances $\overline{R}$.

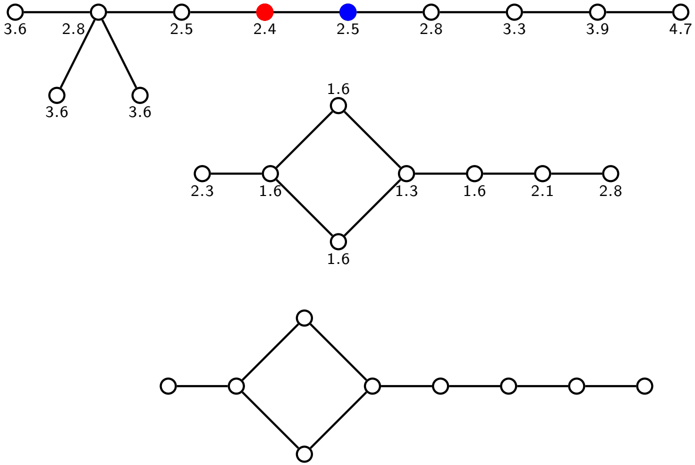






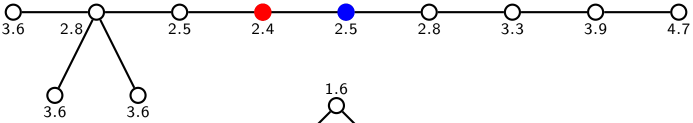
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

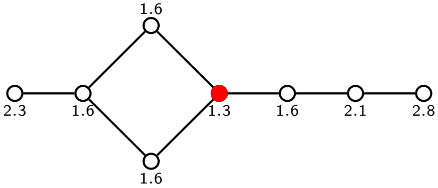
The role of the mean effective resistances $\overline{R}$.

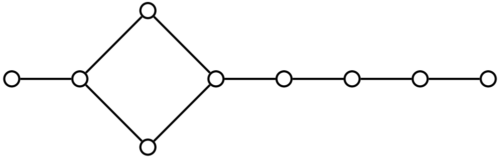


Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

The role of the mean effective resistances $\overline{R}$.

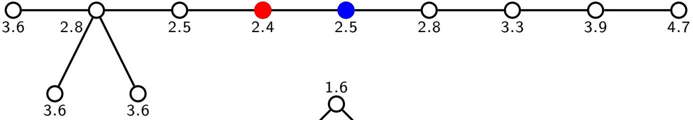


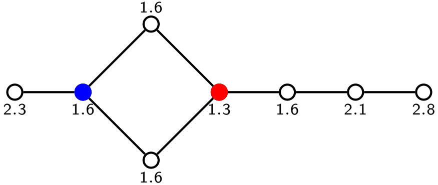


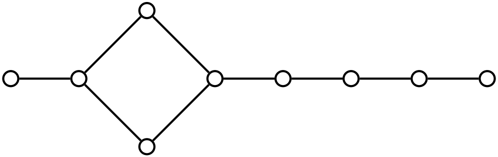


Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

The role of the mean effective resistances $\overline{R}$.

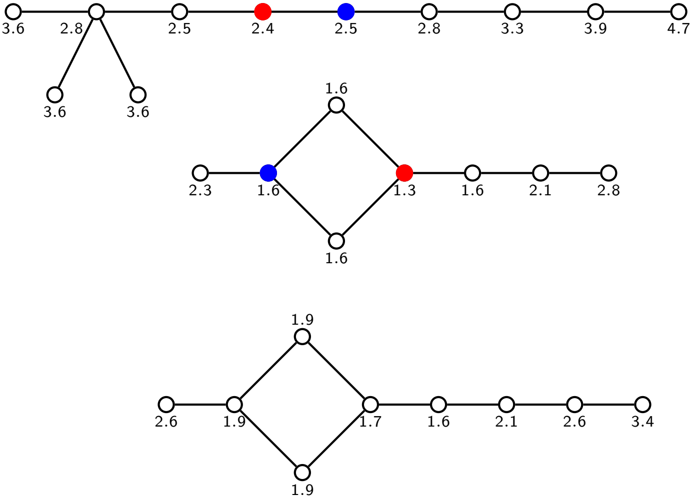






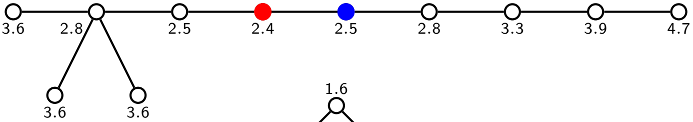
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

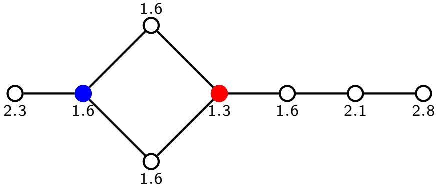
The role of the mean effective resistances $\overline{R}$.

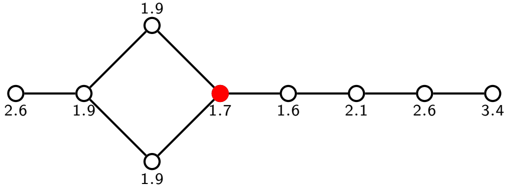


Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

The role of the mean effective resistances $\overline{R}$.

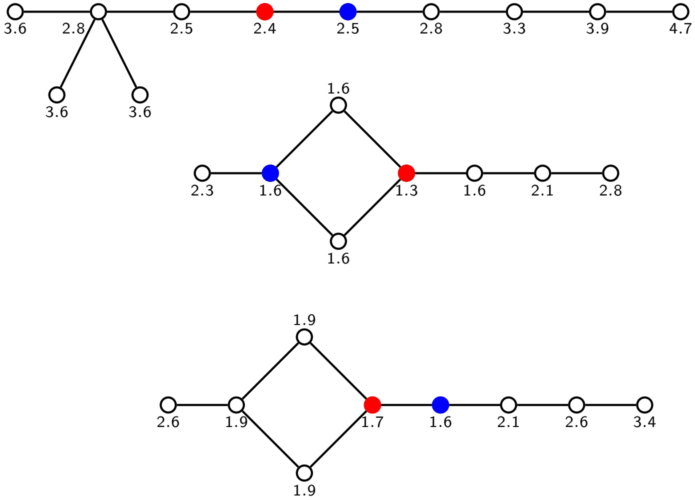






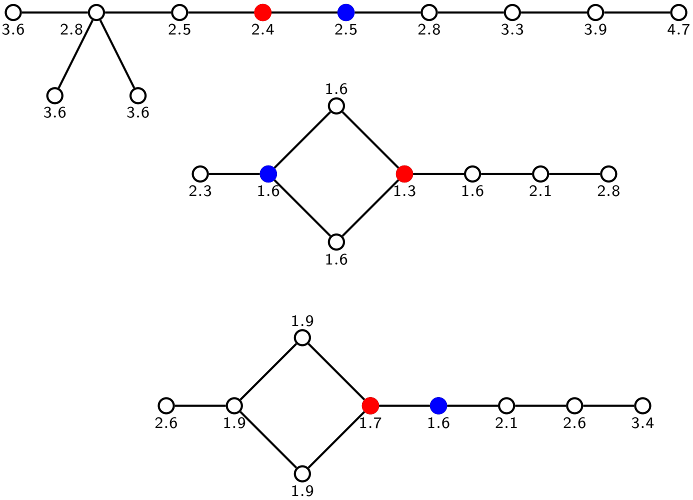
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

The role of the mean effective resistances $\overline{R}$.



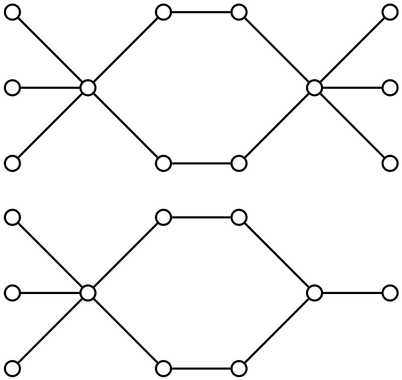
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

The role of the mean effective resistances $\overline{R}$.

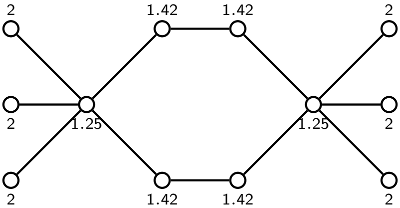


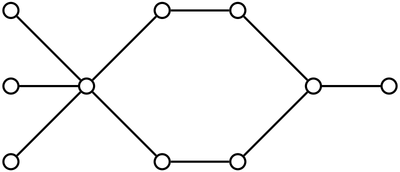
Harmonic centrality $\neq$ Degree centrality

Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

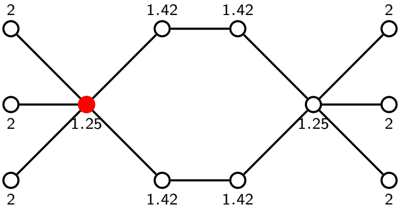


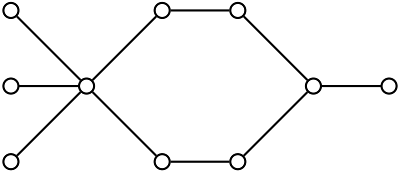
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.



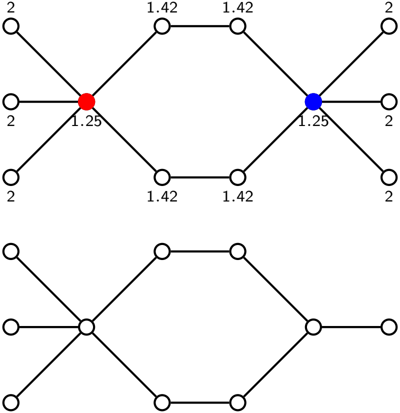


Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1.$

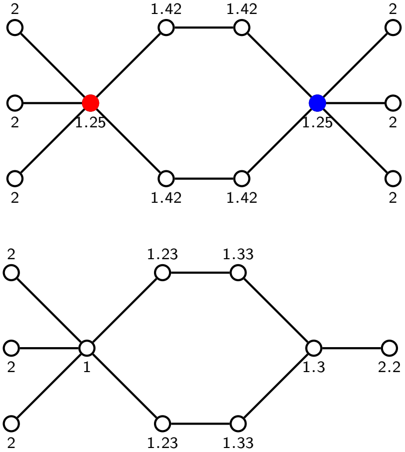




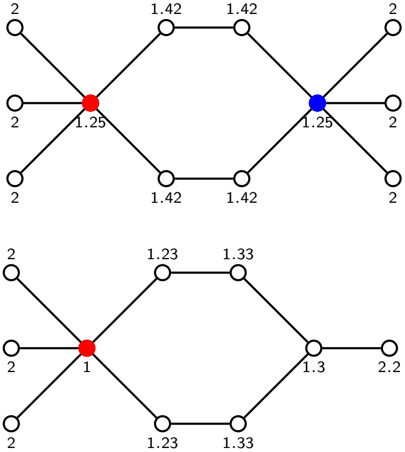
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.



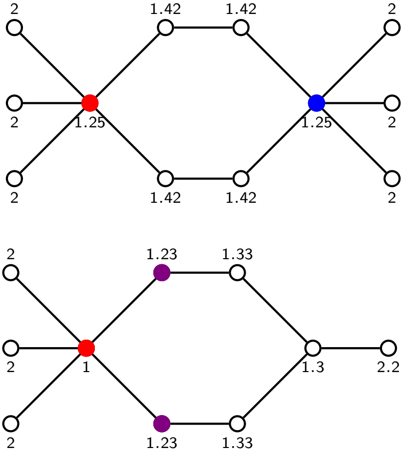
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.



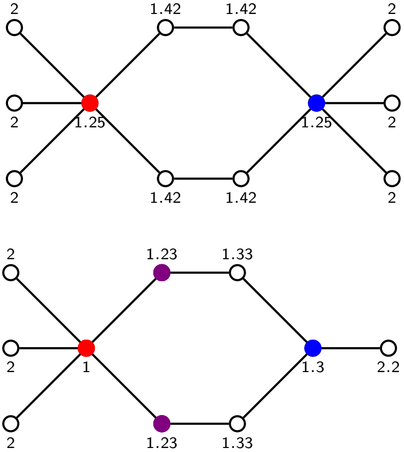
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.



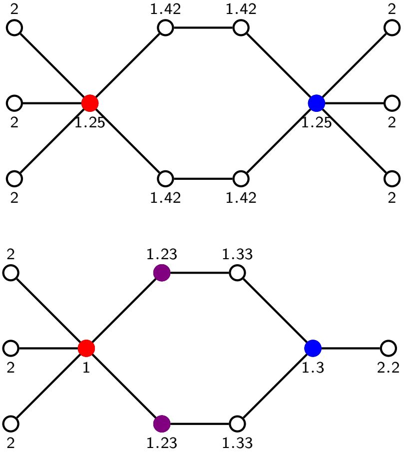
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.



Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

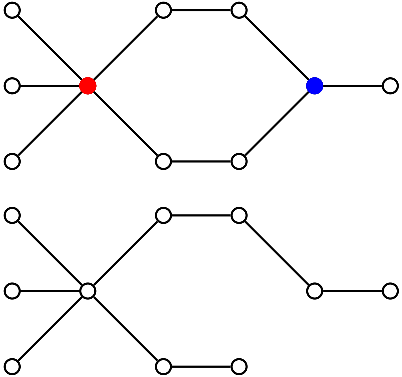


Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

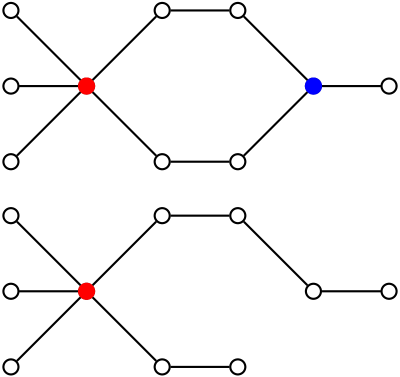


Trade-off: $\min \overline{R}(s_1) \leftrightarrow \max R(s_0 \leftrightarrow s_1)$

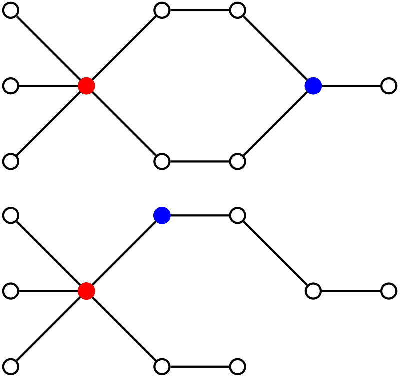
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.



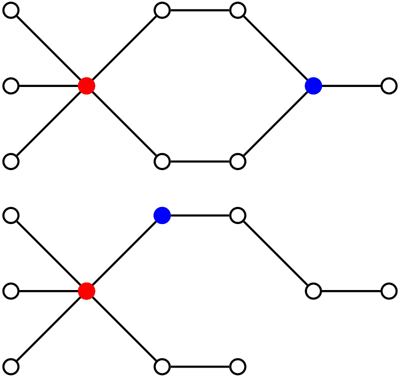
Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.



Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

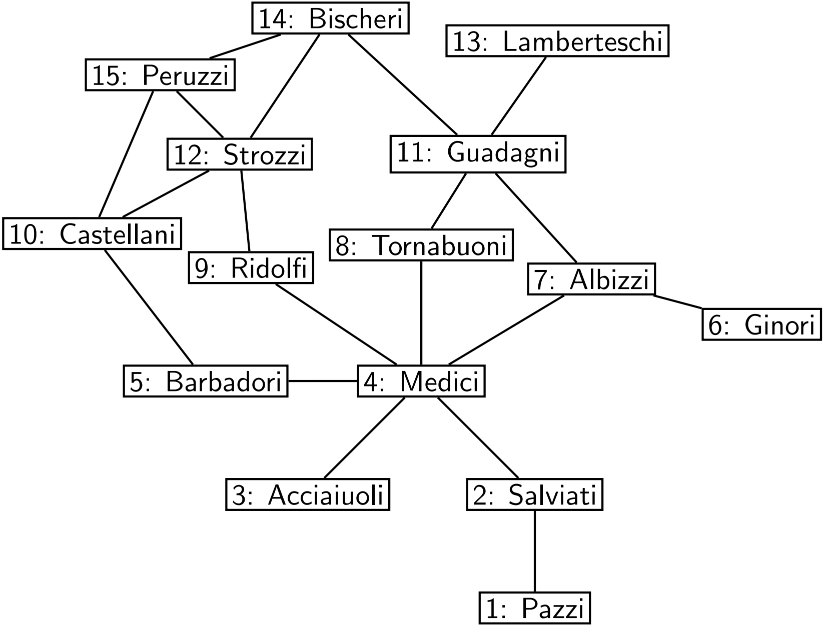


Optimal placement examples, $|\mathcal{S}_0| = |\mathcal{S}_1| = 1$.

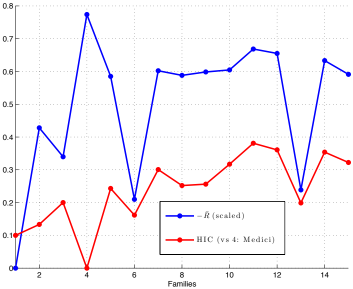


Optimal strategy for s_1: Invade s_0 neighbor or conquer far away virgin areas?

Marriage ties between Florentine families, XV century



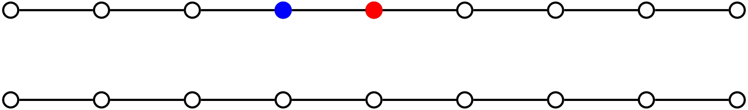
Florentine families: analysis



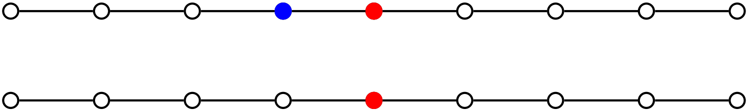
Historical comments:

- Medici (4) is the most authoritative family
- Guadagni (11) is Medici's best opponent
- Strozzi (12), recorded as the main Medici's rival, is only their second best opponent
- the politically weakest Pazzi (1) unsuccessfully attempted an armed conspiracy in 1478

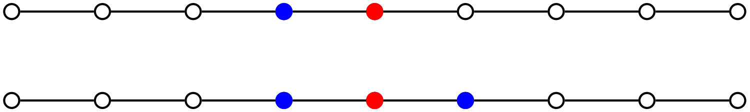
Optimal placement: more general examples



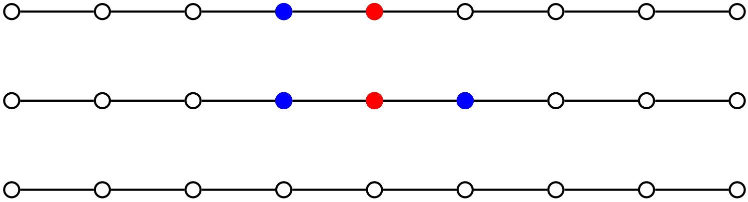
Optimal placement: more general examples



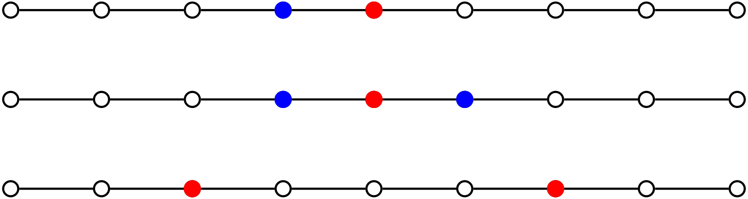
Optimal placement: more general examples



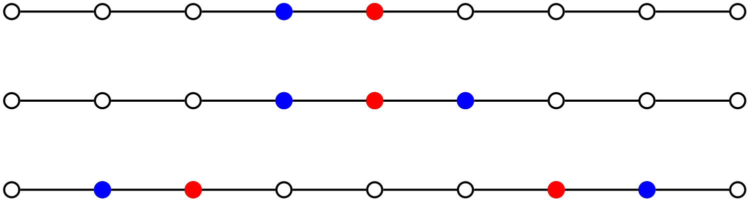
Optimal placement: more general examples



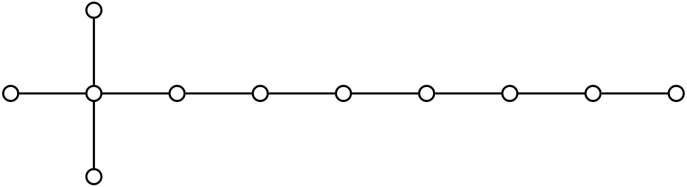
Optimal placement: more general examples



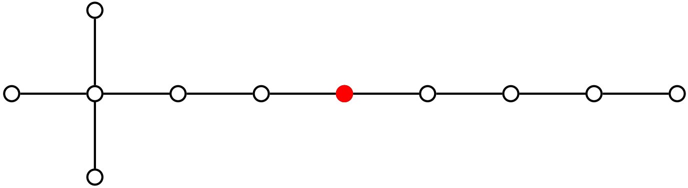
Optimal placement: more general examples



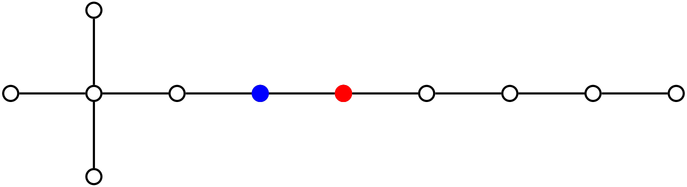
Optimal placement: more general examples



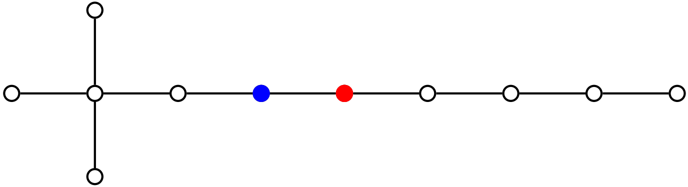
Optimal placement: more general examples

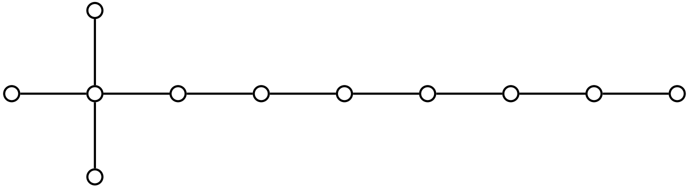


Optimal placement: more general examples

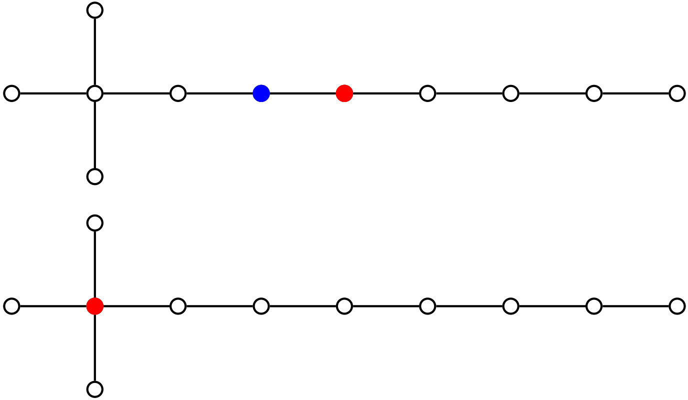


Optimal placement: more general examples

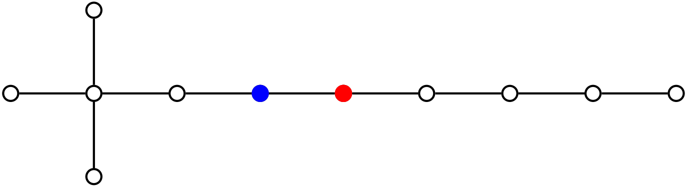


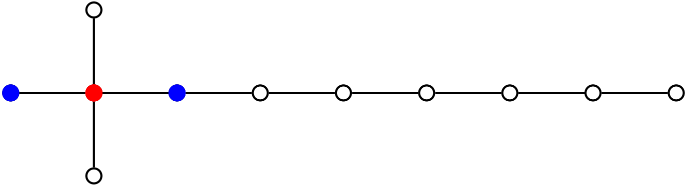


Optimal placement: more general examples

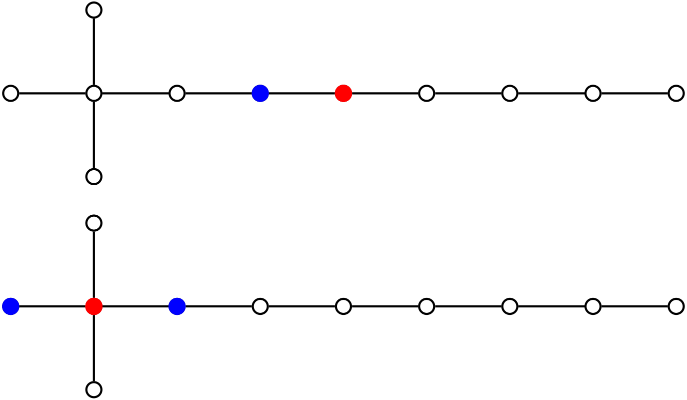


Optimal placement: more general examples





Optimal placement: more general examples



Quite different strategies depending on the number of possible invasors expect!

A distributed message-passing algorithm (the tree case)

- $\mathcal{S}_0$ arbitrary consisting of leaves (no loss of generality).
- $\mathcal{S}_1 = \{\alpha\}$ to be placed. Notation: $H(\alpha)$
- Voltage as a function of α:

$$x_j^{(\alpha)} = \frac{R^{<j\alpha}(\mathcal{S}_0 \leftrightarrow \alpha)}{R^{<j\alpha}(\mathcal{S}_0 \leftrightarrow \alpha) + R(j \leftrightarrow \alpha)}$$

- Voltage scaling $x_j^{(\alpha)} = x_\beta^{(\alpha)} x_j^{(\beta)}$
- A recursive computation for H. $H^{<\alpha\alpha'}$ HIC on $\mathcal{T}^{<\alpha\alpha'}$.

$$(\alpha, \alpha') \in \mathcal{E}, \quad H^{<\alpha\alpha'}(\alpha) = \sum_{\beta \in N_\alpha \setminus \{\alpha'\}} x_\beta^{(\alpha)} H^{<\beta\alpha}(\beta) + 1$$

$$H(\alpha) = \sum_{\beta \in N_\alpha} x_\beta^{(\alpha)} H^{<\beta\alpha}(\beta) + 1$$

A distributed message-passing algorithm

L. Vassio, et al, "Message Passing Optimization of Harmonic Influence Centrality", IEEE T-CONES, Vol. 1(1), pp 109-120, 2014.

Messages are:

- estimate of voltage in α' induced by α: $x^{\alpha\to\alpha'}(t)$
- estimate of influence of α "behind it": $H^{\alpha\to\alpha'}(t)$

For regular agents:

$$H^{\alpha\to\alpha'}(0) = 1 \qquad x^{\alpha\to\alpha'}(0) = 1$$

$$H^{\alpha\to\alpha'}(t+1) = \sum_{\beta\in N_\alpha\setminus\{j\}} x^{\beta\to\alpha}(t) H^{\beta\to\alpha}(t) + 1$$

$$x^{\alpha\to\alpha'}(t+1) = \left(1 + R_{\alpha\alpha'} \sum_{\beta\in N_\alpha\setminus\{\alpha\}} \frac{1 - x^{\beta\to\alpha}(t)}{R_{\beta\alpha}}\right)^{-1}$$

For stubborn leaders:

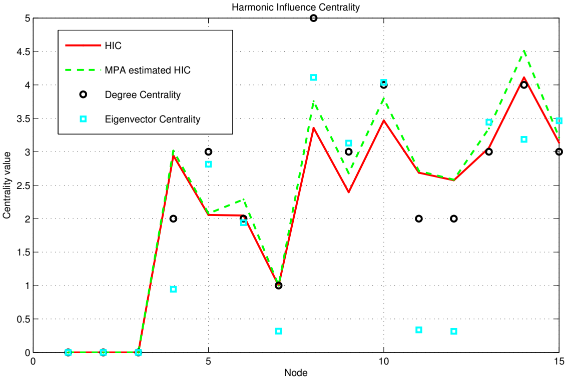
$$H^{s\to\alpha}(t) = 0 \qquad x^{s\to\alpha}(t) = 0 \qquad \text{for all } t$$

Properties of the algorithm.

- **distributed**: can be run by the agents, communicating with neighbors
- **fast**: On trees the message-passing algorithm exactly computes H in $O(\mathrm{diam}(\mathcal{T}))$ steps.
- **low complexity**: On trees (with bounded degrees) the message-passing algorithm exactly computes H in $O(n)$ steps.
- **theoretical analysis** On connected regular graphs, the algorithm converges.
- **simulation** show general remarkable performance

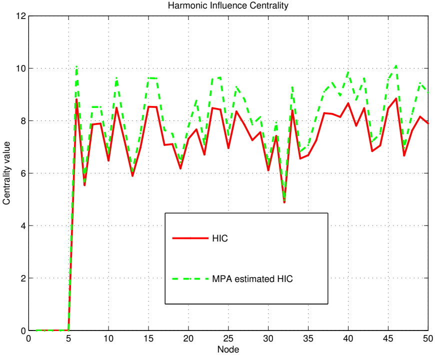
Simulation examples

Erdos-Renyi random graph



Simulation examples

Random 4-regular graph



Harmonic influence in large-scale networks:

Analysis, optimization, and applications from opinion dynamics to distributed estimation

Giacomo Como[1], Fabio Fagnani[2], and **Paolo Frasca**[3]

MTNS 2014
Groningen, July 7, 2014

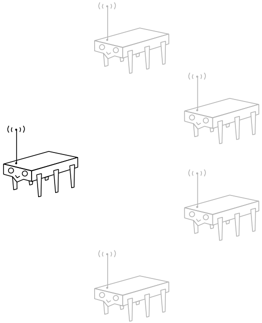
[1]Dept. of Automatic Control, Lund University
[2]Dept. of Mathematical Sciences, Politecnico di Torino
[3]Dept. of Applied Mathematics, University of Twente

Part 3: Estimation from relative measurements

Outline of Part 3

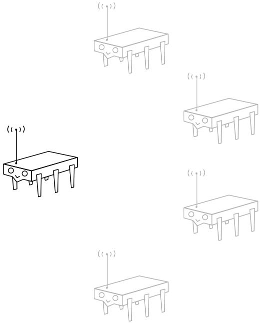
- Estimation from relative measurements
  - Problem statement
  - Applications
- Electrical analogy
- Estimation error
- Anchors and resistances

Relative measurements



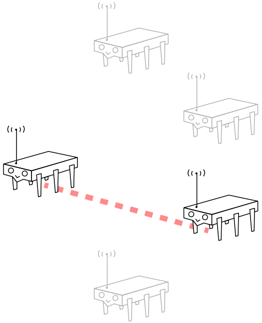
- Sensor $i \in \mathcal{V}$, located at $\bar{x}_i \in \mathbb{R}$

Relative measurements



- Sensor $i \in \mathcal{V}$, located at $\bar{x}_i \in \mathbb{R}$
- $\bar{x}_i \in \mathbb{R}$ is unknown to i

Relative measurements



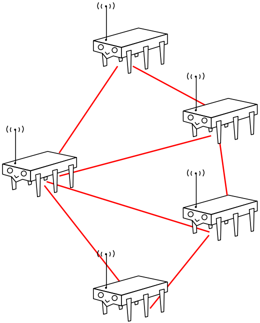
- Sensor $i \in \mathcal{V}$, located at $\bar{x}_i \in \mathbb{R}$
- $\bar{x}_i \in \mathbb{R}$ is unknown to i
- Sensors i and j measure relative values

$$b_{ij} = \bar{x}_i - \bar{x}_j + \nu_{\{i,j\}} \qquad \text{if } i < j$$

with noise $\nu_{\{i,j\}} \sim N(0, \sigma_{ij}^2)$

by symmetry $b_{ji} = -b_{ij}$

Relative measurements



- Sensor $i \in \mathcal{V}$, located at $\bar{x}_i \in \mathbb{R}$
- $\bar{x}_i \in \mathbb{R}$ is unknown to i
- Sensors i and j measure relative values

$$b_{ij} = \bar{x}_i - \bar{x}_j + \nu_{\{i,j\}} \qquad \text{if } i < j$$

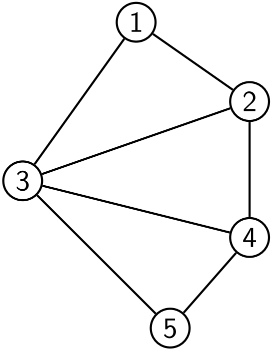
with noise $\nu_{\{i,j\}} \sim N(0, \sigma_{ij}^2)$

by symmetry $b_{ji} = -b_{ij}$

Goal:

Each sensor i seeks an estimate x_i of its actual value $\bar{x}_i$

Estimation from measurements



- $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ connected graph
- define: edge $(i, j) \longleftrightarrow b_{ij}$
- edge (i, j) has weight $\frac{1}{\sigma_{ij}^2}$

Estimation criterion:

$$x = \operatorname*{argmin}_y \frac{1}{2} \sum_i \sum_j \frac{1}{\sigma_{ij}^2} (y_i - y_j - b_{ij})^2$$

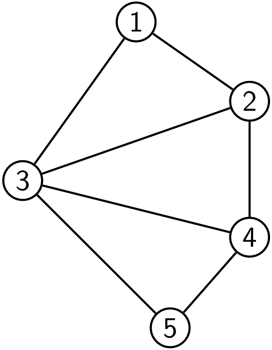
least squares problem, maximum-likelihood estimator

Why this problem?

Applications:

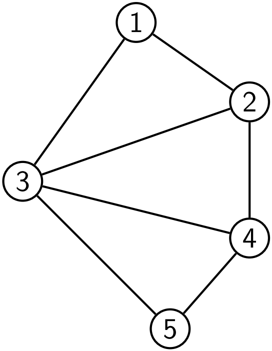
- spatial localization
  - in one dimension, e.g. car platoons
  - in two dimensions (provided the sensors have compasses)
- time synchronization
  by exchanging hello messages, two clocks can measure $t_i - t_j$
- ranking problems $\longrightarrow$ Big Data!
  derive an universal rating from pairwise comparisons
  e.g., Yahoo! and Netflix movie ratings

Electrical interpretation



- $C_{ij} = \frac{1}{\sigma_{ij}^2}$ conductances
- $x = \underset{y}{\operatorname{argmin}} \frac{1}{2} \sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2$

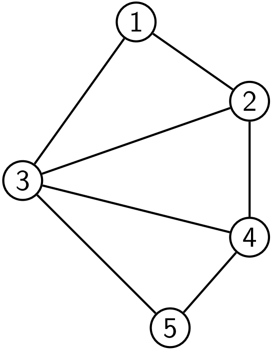
Electrical interpretation



- $C_{ij} = \frac{1}{\sigma_{ij}^2}$ conductances
- $x = \underset{y}{\operatorname{argmin}} \frac{1}{2} \sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2$

$$\frac{1}{2} \sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2$$

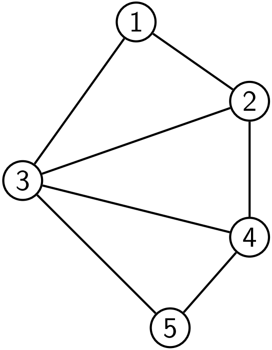
Electrical interpretation



- $C_{ij} = \frac{1}{\sigma_{ij}^2}$ conductances
- $x = \operatorname*{argmin}_y \frac{1}{2} \sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2$

$$\frac{1}{2} \sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2 = \frac{1}{2} \sum_{i,j} C_{ij}(y_i - y_j)^2 + \cancel{\frac{1}{2} \sum_{i,j} C_{ij} b_{ij}^2} - 2 \sum_i y_i \sum_j C_{ij} b_{ij}$$

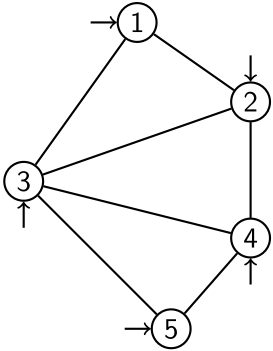
Electrical interpretation



- $C_{ij} = \frac{1}{\sigma_{ij}^2}$ conductances
- $x = \operatorname*{argmin}_y \frac{1}{2} \sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2$
- injected currents $\gamma_i = -\sum_j C_{ij} b_{ij}$
  $\sum_i \gamma_i = 0$

$$\frac{1}{2} \sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2 = \frac{1}{2} \sum_{i,j} C_{ij}(y_i - y_j)^2 + \cancel{\frac{1}{2} \sum_{i,j} C_{ij} b_{ij}^2} - 2 \sum_i y_i \sum_j C_{ij} b_{ij}$$

Electrical interpretation

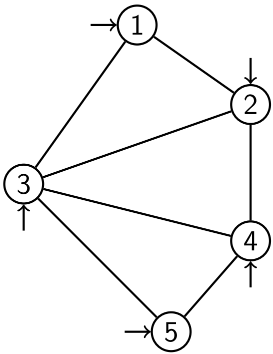


- $C_{ij} = \frac{1}{\sigma_{ij}^2}$ conductances
- $x = \underset{y}{\operatorname{argmin}} \dfrac{1}{2} \displaystyle\sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2$
- injected currents $\gamma_i = -\sum_j C_{ij} b_{ij}$
  $\sum_i \gamma_i = 0$

$$\frac{1}{2}\sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2 = \frac{1}{2}\sum_{i,j} C_{ij}(y_i - y_j)^2 + \cancel{\frac{1}{2}\sum_{i,j} C_{ij} b_{ij}^2} - 2\sum_i y_i \sum_j C_{ij} b_{ij}$$

$$= \underbrace{\frac{1}{2}\sum_{i,j} C_{ij}(y_i - y_j)^2}_{} + 2 \underbrace{\sum_i y_i \gamma_i}_{}$$

Electrical interpretation



- $C_{ij} = \frac{1}{\sigma_{ij}^2}$ conductances
- $x = \underset{y}{\operatorname{argmin}} \frac{1}{2} \sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2$
- injected currents $\gamma_i = -\sum_j C_{ij} b_{ij}$
  $\sum_i \gamma_i = 0$
- minimize energy dissipation with injected currents

$$\frac{1}{2} \sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2 = \frac{1}{2} \sum_{i,j} C_{ij}(y_i - y_j)^2 + \cancel{\frac{1}{2} \sum_{i,j} C_{ij} b_{ij}^2} - 2 \sum_i y_i \sum_j C_{ij} b_{ij}$$

$$= \underbrace{\frac{1}{2} \sum_{i,j} C_{ij}(y_i - y_j)^2}_{\text{energy dissipation}} + 2 \underbrace{\sum_i y_i \gamma_i}_{\text{energy dissipation}}$$

Least squares problem: Solution

Assume: $\sigma_{ij}^2 = 1$ for all i, j

Least squares problem: Solution

Assume: $\sigma_{ij}^2 = 1$ for all i, j

$$x = \underset{y}{\operatorname{argmin}} \frac{1}{2} \sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2$$

satisfies $\sum_j (x_i - x_j) = \sum_j b_{ij}$ that is $Lx = \gamma$

Laplacian equation

Least squares problem: Solution

Assume: $\sigma_{ij}^2 = 1$ for all i, j

$$x = \operatorname*{argmin}_y \frac{1}{2} \sum_i \sum_j C_{ij}(y_i - y_j - b_{ij})^2$$

satisfies $\sum_j (x_i - x_j) = \sum_j b_{ij}$ that is $Lx = \gamma$

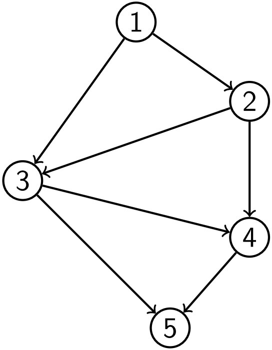
Laplacian equation

We can "invert" the Laplacian using the Green matrix Z:

$$ZL = LZ = I - n^{-1}\mathbf{1}\mathbf{1}', \quad Z_C\mathbf{1} = 0$$

where $\mathbf{1}$ is vector of ones

Least squares problem: Estimation error



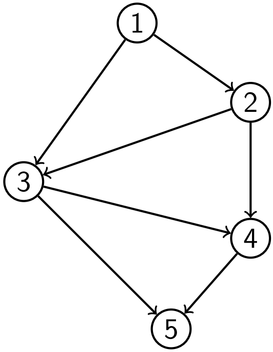
Useful notation:

- edge orientation, incidence matrix: $B \in \mathbb{R}^{\mathcal{E} \times \mathcal{V}} : \; B_{e,i} = -1 \; , \; B_{e,j} = +1$
- vector of measurements $b = B\bar{x} + \nu$

Then, $\gamma = B'b$ and we have

solution: $x = ZB'b + c\mathbf{1}$

Least squares problem: Estimation error



Useful notation:

- edge orientation, incidence matrix: $B \in \mathbb{R}^{\mathcal{E} \times \mathcal{V}}$: $B_{e,i} = -1$, $B_{e,j} = +1$
- vector of measurements $b = B\bar{x} + \nu$

Then, $\gamma = B'b$ and we have

solution: $x = ZB'b + c\mathbf{1}$

Estimation error: averaged over the measurements

$$\frac{1}{n}\mathbb{E}\|x - \bar{x}\|^2$$

or more precisely $$\frac{1}{n}\min_{c \in \mathbb{R}} \mathbb{E}\|(ZB'b + c\mathbf{1}) - \bar{x}\|_2^2$$

Estimation error: "tedious" algebra

$$
\begin{aligned}
\frac{1}{n}\mathbb{E}\|x-\bar{x}\|^2 &= \frac{1}{n}\min_{c\in\mathbb{R}}\mathbb{E}||(ZB'b+c\mathbf{1})-\bar{x}\|_2^2 \\
&= \frac{1}{n}\min_{c\in\mathbb{R}}\mathbb{E}||ZL\bar{x}-\bar{x}+c\mathbf{1}+ZB'\nu||_2^2 \\
&= \frac{1}{n}\min_{c\in\mathbb{R}}\mathbb{E}||\mathbf{1}(-n^{-1}\mathbf{1}'\bar{x}+c)+ZB'\nu||_2^2 \\
&= \frac{1}{n}\min_{c\in\mathbb{R}}\mathbb{E}\left[n(-n^{-1}\mathbf{1}'\bar{x}+c)+\nu'BZ^2B'\nu\right] \\
\text{(choose } c=n^{-1}\mathbf{1}'\bar{x}\text{)} \quad &= \frac{1}{n}\mathbb{E}[\nu'BZ^2B'\nu] \\
&= \frac{1}{n}\mathbb{E}\,\mathrm{tr}[ZB'\nu\nu'BZ] = \frac{1}{n}\mathrm{tr}[ZB'\mathbb{E}[\nu\nu']BZ] = \frac{1}{n}\mathrm{tr}(Z)
\end{aligned}
$$

by Green matrix property $R(i\leftrightarrow j)=Z_{ii}+Z_{jj}-2Z_{ij}$

$$
\implies \quad \frac{1}{n}\mathrm{tr}(Z)=\frac{1}{n^2}\sum_{i,j}R(i\leftrightarrow j)
$$

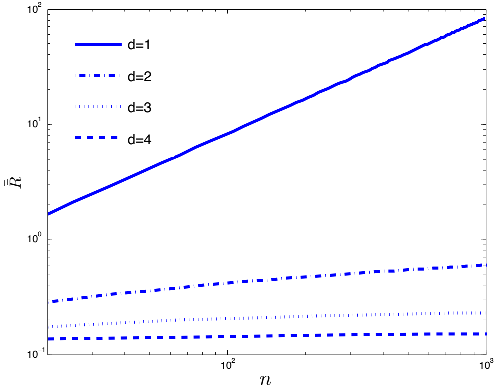
Dimension matters

$$\text{Average resistance } \overline{\overline{R}} = \frac{1}{n^2} \sum_{i,j} R(i \leftrightarrow j)$$

describes how well connected is the network

For d-dimensional graphs:

$$\overline{\overline{R}} \sim \begin{cases} C_1 n & \text{for } d = 1 \\ C_2 \log n & \text{for } d = 2 \\ C_d & \text{for } d \geq 3 \end{cases}$$



High graph dimension gives good performance

Dimension matters

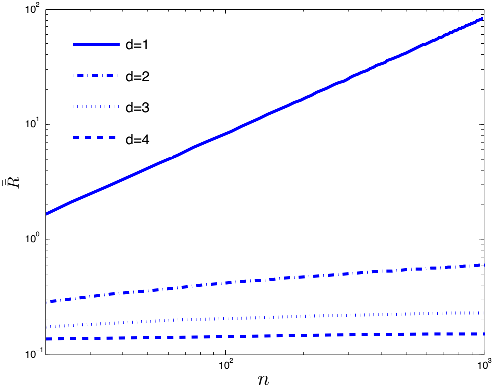
Average resistance $\overline{\overline{R}} = \dfrac{1}{n^2} \sum_{i,j} R(i \leftrightarrow j)$

describes how well connected is the network

For d-dimensional graphs:

$$\overline{\overline{R}} \sim \begin{cases} C_1 n & \text{for } d = 1 \\ C_2 \log n & \text{for } d = 2 \\ C_d & \text{for } d \geq 3 \end{cases}$$

$C_d \sim \frac{1}{d}$ decreasing in d

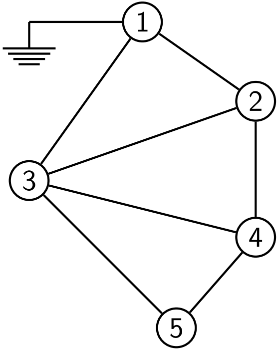


High graph dimension gives good performance

Anchor nodes

If *anchor* node i_0 knows its value exactly,

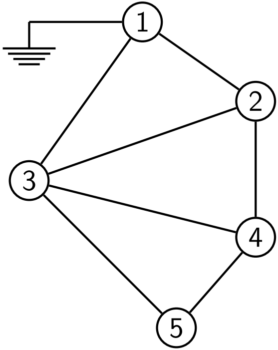
$$x = \operatorname*{argmin}_{y : y_{i_0} = \bar{x}_{i_0}} \frac{1}{2} \sum_i \sum_j \frac{1}{\sigma_{ij}^2} (y_i - y_j - b_{ij})^2$$



Anchor nodes

If *anchor* node i_0 knows its value exactly,

$$x = \operatorname*{argmin}_{y : y_{i_0} = \bar{x}_{i_0}} \frac{1}{2} \sum_i \sum_j \frac{1}{\sigma_{ij}^2} (y_i - y_j - b_{ij})^2$$



Error on sensor i:

$\mathbb{E}|x_i - \bar{x}_i|^2 = R(i_0 \leftrightarrow i)$

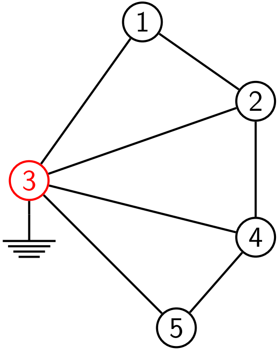
Global error:

$$\frac{1}{n}\mathbb{E}\|x - \bar{x}\|_2^2 = \frac{1}{n}\sum_i R(i_0 \leftrightarrow i) = \overline{R}(i_0)$$

Anchor nodes

If *anchor* node i_0 knows its value exactly,

$$x = \operatorname*{argmin}_{y : y_{i_0} = \bar{x}_{i_0}} \frac{1}{2} \sum_i \sum_j \frac{1}{\sigma_{ij}^2} (y_i - y_j - b_{ij})^2$$



Error on sensor i:

$\mathbb{E}|x_i - \bar{x}_i|^2 = R(i_0 \leftrightarrow i)$

Global error:

$\frac{1}{n}\mathbb{E}\|x - \bar{x}\|_2^2 = \frac{1}{n}\sum_i R(i_0 \leftrightarrow i) = \overline{R}(i_0)$

Conclusion: optimal anchor position is $i_0^\star = \operatorname*{argmin}_j \overline{R}(j)$

Future directions and applications

1. Estimation:
   Design fast, distributed, robust
   algorithms solving the relative estimation problem
   - Can we exploit the electrical analogy
     (as done for harmonic centrality)?

Future directions and applications

1. Estimation:
   Design fast, distributed, robust
   algorithms solving the relative estimation problem
   - Can we exploit the electrical analogy
     (as done for harmonic centrality)?
2. Experimental design:
   optimize anchor position, overall topology, addition of edges
   - Can we find distributed algorithms?

References

- P. Barooah and J. P. Hespanha. Estimation from relative measurements: Algorithms and scaling laws. *IEEE Control Systems Magazine*, 27(4):57–74, 2007
- A. Giridhar and P. R. Kumar. Distributed clock synchronization over wireless networks: Algorithms and analysis. In *IEEE Conference on Decision and Control*, pages 4915–4920, San Diego, CA, USA, December 2006
- W. S. Rossi, P. Frasca, and F. Fagnani. Effective resistance of toroidal graphs; some sharper results and applications, September 2013. Submitted for publication
- B. Osting, C. Brune, and S. Osher. Optimal data collection for improved rankings expose well-connected graphs. *Journal of Machine Learning*, 2014. to appear