From local averaging dynamics to the emergence of global phenomena: the fundamental role of network interconnection **(special issue JCW)**

Giacomo Como and Fabio Fagnani *

This work is dedicated to the living memory of Jan C. Willems.

Abstract

Averaging dynamics are among the simplest and most studied network dynamics with a large variety of applications ranging from cooperative inferential algorithms over sensor networks, to robot formation, and opinion dynamics models. They can be seen as the result of interconnecting simple linear devices much in the style of the abstract concept of interconnection proposed by Jan Willens in its behavioral approach to systems and control. In this paper, we gather and originally present a number of fundamental results and examples scattered in the literature showing the neat and deep interplay existing between the topology of the network and the resulting global emerging behavior in the system.

1 Introduction

One of the core concepts in the behavioral approach to systems and control developed by Jan Willems in the '80's is that of an interconnection [9]. Encompassing the traditional feedback interconnection on which classical input/output control theory was based, the behavioral approach allows for defining interconnection of systems at a more primitive level, as intersection of solution sets of the evolution equations, without the need for considering specific flow diagrams. As Jan used to repeat many times 'what is an input and what is an output is a matter of the application you have in mind'. This idea of going beyond classical input/output formalism was proven to be very fruitful in the applications to coding theory and, in particular, his study of the minimal state space realizations [8] lays at the foundation of trellis representations which are the basic tool for the design of efficient decoding algorithms.

More recently, network dynamics is also featuring deep cultural connection with the *ansatz* of the behavioral approach to systems. Network dynamics entail a large number of (typically) simple dynamical systems coupled together along the architecture of a graph and the overall dynamical system can thus be seen as the interconnection of these simple atomic devices. It does not make much sense to classify a-priori interconnection signals

*G. Como is with the Department of Automatic Control, Lund University, SE-221 00 Lund, Sweden giacomo.como@control.lth.se. F. Fagnani is with the Giuseppe Luigi Lagrange Department of Mathematical Sciences, Politecnico di Torino, Corso Stati Uniti 24, 10129, Torino, Italy, fabio.fagnani@polito.it.

in input and outputs, rather they are variables coupling the systems, possibly sensor measurements, state positions, epidemic states, and it is often impossible to say who is influencing who. The emergence of global behavior like coordination, diffusive and fusion phenomena is one of the most distinctive features of these interconnected systems. The global behavior can be seen as the result of the local behaviors and of the topology of the graph which determines the architecture of the interconnection.

For the sake of describing a particularly simple and very well studied case of network dynamics, in these paper we will focus on distributed averaging systems: this is a linear network dynamics already exhibiting a number of interesting features of collective behaviors, like synchronization and phase transition phenomena. Its applications range from inferential sensor network algorithms, to network robot formation to models for opinion dynamics. Most of the behavioral approach developed by Jan was indeed focused on linear systems theory: he used to say that the theory was already sufficiently rich at that level from a theoretical point of view and already containing such an enormous set of different applications. Keep things as simple as possible was a fundamental requirement in Jan's approach to science.

Using classical results from Perron-Frobenius theory of non-negative matrices, we will first present an asymptotic analysis of a linear averaging dynamics on arbitrary interconnection graphs. As expected, the topology of the graph plays a crucial role in the behavior. While for connected graph, it is well known that all states reach an asymptotic consensus, Theorem 2 analyze the case of a general graph and shows, in particular, that the asymptotic state of any agent in the network turns out to be a convex combination of the consensus state reached by the sink connected components (components from which there is no outgoing edge). The weights of such convex combination have a series of interesting and useful interpretations. They can be seen as hitting probabilities of the dual Markov process generated by the same averaging matrix or, in case when the graph is symmetric, as voltages of an electrical circuit with a suitable boundary condition on the nodes belonging to the sink components. An especially relevant case in the applications is when the sink components are all constituted by singleton nodes. Such nodes can be interpreted as stubborn nodes never changing their state, for instance playing the role of opinion leaders in social networks, or leaders in formation control. The final part of this paper is dedicated to a deeper understanding of how the asymptotic state is distributed in the network in the presence of such stubborn nodes. It turns out that depending on the strength of the stubborn nodes and the connectivity of the graph quite different phenomena can show up passing from 'polarized' to 'homogeneous' equilibrium configurations. In the polarized case nodes tends to split in subfamilies and get the opinion of a particular stubborn agent, while in the homogeneous regime most of the nodes tend to get close to a consensus on a value which depends on the state of the various stubborn nodes. In this paper we will present these phenomena through an example where the transition phase between the two regimes can be analyzed in detail. We will then recall more abstract results already appeared in the literature.

The intrinsic simplicity of the linear averaging network dynamics and, at the same time, this beautiful and neat interplay between the graph describing the architecture of the interconnections, and the structure of the solutions, we believe, lays in that part of science that Jan mostly appreciated: simple and deep at the same time.

1.1 Notation

For the convenience of the reader, we gather below a number of conventional notation used throughout the paper. A' is the transpose of a matrix A; $\mathbb{1}$ is the all-1 vector; $\delta^{(i)}$ is the vector with all entries equal to 0 except for the i-th that is equal to 1; for a vector x and a subset of indices $\mathcal{A}$, $x_{\mathcal{A}} := \sum_{i\in\mathcal{A}} x_i$, while $x_* = \min_i x_i$ is the minimum entry of x.

2 Averaging dynamics on arbitrary graphs

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ be a directed weighted graph representing the network, where $\mathcal{V} = \{1, \dots, n\}$ is the set of nodes, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of links, and $W \in \mathbb{R}^{n\times n}$ is a matrix of nonnegative link weights such that $W_{ij} > 0$ if and only if $(i, j) \in \mathcal{E}$, with positive diagonal elements of W corresponding to self-loops. We will refer to the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ as: (strongly) *connected* if W is irreducible (e.g. no permutation can transform it into a block diagonal matrix); *undirected* if W is symmetric; *balanced* if $W\mathbb{1} = W'\mathbb{1}$; *unweighted* if $W_{ij} \in \{0, 1\}$ for all $i, j \in \mathcal{V}$. We will denote the out-degree vector by $w = W\mathbb{1}$ and assume[1] that $w_i > 0$ for all nodes i. We then associate to the graph $\mathcal{G}$ the matrices

$$D = \mathrm{diag}\,(w)\,, \qquad P = D^{-1}W\,. \tag{1}$$

Observe that P is a stochastic matrix, i.e., it is nonnegative and such that $P\mathbb{1} = \mathbb{1}$. Also, note that $P'w = w$ if and only if $\mathcal{G}$ is balanced. Moreover, the property that $\mathcal{G}$ is undirected is equivalent to the detailed balance equation $w_iP_{ij} = w_jP_{ji}$ for all $i, j \in \mathcal{V}$, a property that is referred to as reversibility of P (with respect to w).

One of the most popular network dynamical systems can be seen as the interconnection of locally averaging systems. A local averaging system placed at node $i \in \mathcal{V}$ is a multi-input/single-state dynamics governed by the linear update rule $x_i(t + 1) = \alpha x_i(t) + (1 - \alpha) \sum_{j\in\mathcal{V}} P_{ij}u_j(t)$. Here $\alpha \in [0, 1]$ is an inertia parameter. By putting $u_j(t) = x_j(t)$ for all $j \in \mathcal{V}$, one obtains the interconnected system

$$x_i(t + 1) = \alpha x_i(t) + (1 - \alpha) \sum_{j\in\mathcal{V}} P_{ij}x_j(t), \qquad i \in \mathcal{V}\,. \tag{2}$$

In the above, the summation index j runs in principle over the whole node set $\mathcal{V}$, but is in fact restricted to the out-neighborhood $\mathcal{N}_i := \{j \in \mathcal{V} : W_{ij} > 0\}$ of node i in $\mathcal{G}$. Assembling all the nodes' states in a column vector $x(t) \in \mathbb{R}^n$, we can rewrite (2) in the compact form

$$x(t + 1) = P_\alpha x(t) \tag{3}$$

where

$$P_\alpha = \alpha I + (1 - \alpha)P\,. \tag{4}$$

Hence, the state vector $x(t)$ of the distributed averaging dynamics (3) evolves as

$$x(t) = P_\alpha^t x(0)\,, \qquad t \geq 0\,, \tag{5}$$

so that its asymptotic behavior is dictated by the eigen-structure of P_α. In particular, since P is a stochastic matrix, it is non-expansive in the $||\cdot||_\infty$ norm, so that its spectrum

[1] This assumption implies no loss of generality since one can always add a self-loop with positive weight W_{ii} to nodes i with zero out-degree without modifying connectivity and other properties of the network.

is contained in the unitary disk centered in 0. Hence, for all $0 \leq \alpha \leq 1$, the matrix P_α has one eigenvalue equal to 1 and its whole spectrum is contained in the closed disk of diameter coinciding with the segment joining the points $-1+2\alpha$ and 1 in the complex plane. What is relevant in this context is that finer properties of the spectrum of P_α are closely related to geometrical properties of the graph $\mathcal{G}$ as we briefly summarize below.

First we consider the case when the graph $\mathcal{G}$ is connected. In this case, it is a standard result of Perron-Frobenius theory that P_α^t converges to a matrix $\mathbb{1}\pi'$ where π can be uniquely characterized as the only right eigenvector $\pi = P'\pi$ normalized in such a way that $\mathbb{1}'\pi = 1$. This vector π has all components strictly positive and will be referred to as the *centrality* vector of $\mathcal{G}$ (notice that is a probability vector). For a connected balanced graph, π is proportional to the degree vector, namely, $\pi = w/(\mathbb{1}'w)$. For general, not necessarily balanced, connected graphs such simple expression does not hold true, while one can express the entries π_i in terms of infinite sums. For $\alpha \in [0,1)$, we define the *mixing time* of P_α as

$$\tau_\alpha := \inf \left\{ t \geq 0 : \max_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}} |(P_\alpha^t)_{ij} - \pi_j| \leq \frac{1}{e} \right\}$$

The mixing time is a popular index to study the speed of convergence of P_α^t. In certain cases can be estimated from the knowledge of the second largest eigenvalue of the matrix P_α. This happens for instance when $\mathcal{G}$ is the unweighted d-dimensional toroidal grid: in this case $\tau_\alpha \sim C_d n^{2/d}$, where C_d is a constant depending on the dimension d but independent of the graph size n. For more general large-scale graphs the direct analysis of the spectrum is often unfeasible and in these cases it is useful to relate the mixing time to geometric properties of the graph. Specifically, we consider the so called graph conductance

$$\Phi := \min_{\emptyset \neq \mathcal{U} \subsetneq \mathcal{V}} \frac{\sum_{i \in \mathcal{U}} \sum_{j \in \mathcal{V} \setminus \mathcal{U}} \pi_i P_{ij}}{\sum_{i \in \mathcal{U}} \pi_i \cdot \sum_{j \in \mathcal{V} \setminus \mathcal{U}} \pi_j}$$

which measures the lack of bottlenecks in the graph. Then, results in [6, Section 4.3] imply that

$$\frac{1-2/e}{\Phi} \leq \tau_{1/2} \leq \frac{1}{\Phi^2} \log \frac{e^2}{\pi_*},$$

where $\pi_* = \min_{i \in \mathcal{V}} \pi_i$. Using these techniques, for instance, it can be shown that the popular unweighted Erdos-Renyi random graphs in the connected regime[2] exhibit, with probability 1, mixing times of the order of $\log n$. Such graphs are thus mixing much faster then the ones of the d-dimensional toroidal grids.

Going back to the distributed averaging dynamics (3), previous results imply that the state $x(t)$ always converges to a consensus vector $\overline{x}\mathbb{1}$ on a value $\overline{x} = \pi' x(0)$ that is the weighted average of the the nodes' initial values, with weights corresponding to the entries of the centrality vector. The speed of this convergence is captured by the mixing time τ_α. Precisely, we have the following result.

Proposition 1. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ be a connected graph. Then, for every $\alpha \in (0,1)$, the distributed averaging dynamics* (3) *satisfy*

$$||x(t) - \mathbb{1}\overline{x}||_\infty \leq ||x(0) - \mathbb{1}\overline{x}||_\infty \exp(-\lfloor t/\tau_\alpha \rfloor), \tag{6}$$

[2] they are constructed by randomly putting an edge between any pair of n fixed nodes independently with probability $p = c \log n/n$ with $c > 1$

for all $x(0) \in \mathbb{R}^n$*, where*

$$\overline{x} = \pi^* x(0)\,, \tag{7}$$

so that, in particular,

$$\lim_{t\to+\infty} x(t) = \mathbb{1}\overline{x}\,.$$

We now move on to discussing the asymptotic behavior of the distributed averaging dynamics (3) in arbitrary —not necessarily connected— graphs. The node set $\mathcal{V}$ can always be uniquely partitioned as $\mathcal{V} = \mathcal{V}_1 \cup \cdots \cup \mathcal{V}_c$ where the subgraphs (called connected components) $\mathcal{G}_k = (\mathcal{V}_k, \mathcal{E} \cap (\mathcal{V}_k \times \mathcal{V}_k), W_{|\mathcal{V}_k \times \mathcal{V}_k})$ are connected and maximal with respect to this property. Given two connected components $\mathcal{G}_h$ and $\mathcal{G}_k$, we say that $\mathcal{G}_h \geq \mathcal{G}_k$ if there is a path in $\mathcal{G}$ connecting some node in $\mathcal{V}_k$ to some node in $\mathcal{V}_h$. By construction the relation is transitive and such that $\mathcal{G}_h \geq \mathcal{G}_k$ and $\mathcal{G}_k \geq \mathcal{G}_h$ if and only if $\mathcal{G}_k = \mathcal{G}_h$. Connected components which are maximal with respect to the partial ordering $\geq$ are called sink components: any path starting in a sink component will never leave it.

The following result describes the asymptotic behavior of the averaging dynamics (3) for a general graph.

Theorem 2. *Let* $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ *be a graph with sink components* $\mathcal{G}_k = (\mathcal{S}_k, \mathcal{E} \cap (\mathcal{S}_k \times \mathcal{S}_k), W_{|\mathcal{S}_k \times \mathcal{S}_k})$*, for* $k = 1, \ldots, s$*. Let* $\mathcal{S} := \bigcup_{1\leq k\leq s} \mathcal{S}_k$*. Then, there exists a unique stochastic matrix* $H \in \mathbb{R}^{n\times s}$ *such that, for all* $\alpha \in [0, 1]$*,*

$$(I - P_\alpha)H = 0\,, \qquad H_{ik} = \begin{cases} 1 & \text{if} \quad i \in \mathcal{S}_k \\ 0 & \text{if} \quad i \in \mathcal{S}_{-k} = \mathcal{S} \setminus \mathcal{S}_k \end{cases}\,, \qquad i \in \mathcal{S}\,, \quad 1 \leq k \leq s\,, \tag{8}$$

such that, for every $\alpha \in (0, 1)$*, the distributed averaging dynamics* (3) *satisfies*

$$\lim_{t\to+\infty} x(t) = H\overline{x}\,, \tag{9}$$

for all $x(0) \in \mathbb{R}^n$*, where* $\overline{x} \in \mathbb{R}^s$ *has entries*

$$\overline{x}_k = \sum_{i\in\mathcal{S}_k} \pi_i^{(k)} x_i(0)\,, \qquad k = 1, \ldots, s\,. \tag{10}$$

where $\pi^{(k)} \in \mathbb{R}^{\mathcal{S}_k}$ *is the centrality vector of* $\mathcal{G}_k$*.*

Proof If we order the nodes in such a way that sink connected components come last, for every $\alpha \in (0, 1)$ the stochastic matrix P_α takes the form

$$P_\alpha = \begin{pmatrix} Q & R^{(1)} & \ldots & \ldots & R^{(s)} \\ 0 & P^{(1)} & 0 & \ldots & 0 \\ 0 & 0 & P^{(2)} & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & P^{(s)} \end{pmatrix}$$

Splitting the state vector $x(t) = (y(t), x^{(1)}(t), \ldots, x^{(s)}(t))$ accordingly, with $y(t) \in \mathbb{R}^{\mathcal{R}}$ where $\mathcal{R} := \mathcal{V} \setminus \mathcal{S}$, and $x^{(k)}(t) \in \mathbb{R}^{\mathcal{S}_k}$ for $k = 1, \ldots, s$, the dynamics (3) is reduced to

$$\begin{aligned} y(t+1) &= Qy(t) + \sum_{1\leq k\leq s} R^{(k)} x^{(k)}(t) \\ x^{(k)}(t+1) &= P^{(k)} x^{(k)}(t) \qquad\qquad 1 \leq k \leq s\,. \end{aligned} \tag{11}$$

First, notice that the evolution of the state on the nodes of the sink components can be studied using Proposition 1. In particular, for $k = 1, \dots, s$, we have that $x^{(k)}(t) \stackrel{t\to\infty}{\longrightarrow} \mathbb{1}\overline{x}_k$, where $\overline{x}_k$ is as in (10) and $\pi^{(k)} \in \mathbb{R}^{\mathcal{S}_k}$ is the centrality vector of $\mathcal{G}_k$.

On the other hand, Q is a sub-stochastic matrix, i.e., it is nonnegative and such that $\sum_{j\in\mathcal{R}} Q_{ij} \leq 1$ for all $i \in \mathcal{R}$. By the Perron-Frobenius theory, the spectral radius ρ of Q is an eigenvalue with nonnegative associated right eigenvector $z \in \mathbb{R}^{\mathcal{R}}$. Let $\mathcal{J} \subseteq \mathcal{R}$ is the support of z. Then, since from every node in $\mathcal{R}$ is connected to some node in $\mathcal{S}$, necessarily $\min_{j\in\mathcal{J}} \sum_{i\in\mathcal{J}} Q_{ji} < 1$ (otherwise there would be no links from $\mathcal{J}$ to $\mathcal{V} \setminus \mathcal{J}$). Hence, one has that

$$\rho \sum_{i\in\mathcal{J}} z_i = \sum_{i\in\mathcal{J}} \sum_{j\in\mathcal{J}} Q_{ji} z_j < \sum_{j\in\mathcal{J}} z_j \,,$$

i.e., $\rho < 1$. This implies that $I - Q$ is invertible with nonnegative inverse $(I - Q)^{-1} = \sum_{l\geq 0} Q^l$. Since $x^{(k)}(t) \to \mathbb{1}\overline{x}_j$ for $1 \leq k \leq s$, classical results on linear systems then imply that the vector

$$y(t) = Q^t y(0) + \sum_{1\leq l\leq t} Q^l \sum_{1\leq k\leq s} R^{(k)} x^{(k)}(t-l) \stackrel{t\to\infty}{\longrightarrow} (I-Q)^{-1} \sum_{1\leq k\leq s} R^{(k)} \mathbb{1}\overline{x}_k \,.$$

This yields (9), by letting $H \in \mathbb{R}^{n\times s}$ have entries

$$H_{ik} = \begin{cases} ((I-Q)^{-1} R^{(k)} \mathbb{1})_i & \text{if } i \in \mathcal{R} \\ 1 & \text{if } i \in \mathcal{S}_k \\ 0 & \text{if } i \in \mathcal{S} \setminus \mathcal{S}_k \,. \end{cases} \tag{12}$$

Observe that $P_\alpha \mathbb{1} = \mathbb{1}$ implies $\sum_{1\leq k\leq s} R^{(k)} \mathbb{1} = (I - Q)\mathbb{1}$, so that

$$\sum_{1\leq k\leq s} H_{ik} = ((I-Q)^{-1} \sum_{1\leq k\leq s} R^{(k)} \mathbb{1})_i = 1 \,, \qquad i \in \mathcal{R} \,.$$

Since clearly $\sum_{1\leq k\leq s} H_{ik} = 1$ for all $i \in \mathcal{S}$, we get that $H\mathbb{1} = \mathbb{1}$. Moreover, H is nonnegative since both $(I - Q)^{-1}$ and $R^{(k)}$, $1 \leq k \leq s$, are. Hence H is stochastic. Finally, one has that

$$(I - P_\alpha)H = (I - Q)(I - Q)^{-1} \sum_{1\leq k\leq s} R^{(k)} \mathbb{1} (\delta^{(j)})' - \sum_{1\leq k\leq s} R^{(k)} \mathbb{1} (\delta^{(k)})' = 0 \,,$$

thus showing that (8) holds true. ■

Theorem 2 states that the nodes belonging to a sink components $\mathcal{S}_k$ asymptotically reach consensus on the value $\overline{x}_k = \sum_{i\in\mathcal{S}_k} \pi_i^{(k)} x_i(0)$. Every other node $i \in \mathcal{R}$ has its state convergent to a convex combination of the consensus values of the various sink components (with weights given by the coefficients H_{ik}. The initial state of the nodes $i \in \mathcal{R}$ have thus no effect on the final state.

We analyze two extreme cases fitting in Theorem 2.

- If the number of sink components $s = 1$, then, necessarily, $H = \mathbb{1}$ and Theorem 2 implies that $P_\alpha^t x(0) \stackrel{t\to\infty}{\longrightarrow} \mathbb{1}\pi^* x(0)$. I.e., the system converges to consensus where the consensus point is a convex combination of the initial condition of the nodes of the sink component, while all other nodes do not play any role. This can be seen as a generalization of Proposition 1.

- If the sink components are all singletons: $\mathcal{S}_k = \{v_k\}$ for $k = 1, \dots, s$, then $P_\alpha^t x(0) \stackrel{t\to\infty}{\longrightarrow} H\overline{x}$, where $\overline{x}_k = x_{v_k}(0)$ for $1 \le k \le s$. Sink nodes in this case maintain their state constant in time, while all other nodes converge to equilibrium values that can be expressed as convex combinations of the states of the sink nodes. Sink nodes — sometimes referred to as *stubborn nodes*— are used to model opinion leaders in social networks [1], or, in robotic formation applications, anchor nodes who do not change their position [4].

An interesting probabilistic interpretation comes by considering a discrete-time Markov chain $X(t)$ with state space $\mathcal{V}$ and transition probability matrix P_α, i.e., for all $t \ge 0$, $X(t+1)$ is conditionally independent from the past $X(0), X(1), \dots, X(t-1)$ given the present $X(t)$, and $\mathbb{P}(X(t+1) = j | X(t) = i) = (P_\alpha)_{ij}$. It is a well known fact that, with probability one, $X(t)$ will enter one of the sink components in finite time and never leave it ever after. For $k = 1, \dots, s$, let A_k be the event that $X(t)$ enters the sink component $\mathcal{S}_k$ before any other. Consider the matrix $M \in \mathbb{R}^{n\times s}$ given by $M_{ik} = \mathbb{P}(A_k \,|\, X(0) = i)$. Then, a straightforward conditioning argument yields

$$M_{ik} = \mathbb{P}(A_k \,|\, X(0) = i) = \sum_{l\in\mathcal{V}} (P_\alpha)_{il} \mathbb{P}(A_k \,|\, X(1) = l) = (P_\alpha M)_{ik} \tag{13}$$

for all $i \in \mathcal{V}$ and $k = 1, \dots, s$. Hence, M solves (8) and we can deduce that $H = L$. In other terms, the coefficient H_{ik} which is the weight that agent i puts on $\overline{x}_k$ in determining its asymptotic state x_i, can be interpreted as the hitting probability that a Markov chain started at node i and moving with transition probability matrix P_α hits the sink component $\mathcal{S}_k$ before any other sink component.

3 Electrical network interpretation

Consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ with sink components $\mathcal{S}_1, \dots, \mathcal{S}_s$, where $s \ge 2$. Put $\mathcal{S} := \bigcup_{1\le k\le s} \mathcal{S}_k$, $\mathcal{R} := \mathcal{V} \setminus \mathcal{S}$, and $\mathcal{S}_{-k} := \mathcal{S} \setminus \mathcal{S}_k$ for $1 \le k \le s$. Then, we know from Theorem 2 that the distributed averaging dynamics (3) converges to an equilibrium vector $x = H\overline{x}$, where $H \in \mathbb{R}^{n\times s}$ is the stochastic matrix satisfying (8) and $\overline{x} \in \mathbb{R}^s$ is the vector of the weighted averages of the initial condition in the s sink components.

In this section, we focus on the special case when the restriction of the graph $\mathcal{G}$ to $\mathcal{R}$ is connected and undirected, i.e., when

$$W_{ij} = W_{ji}\,, \qquad i, j \in \mathcal{R}\,. \tag{14}$$

Define the link flows

$$f_{ij} = W_{ij}(x_i - x_j)\,, \quad i, j \in \mathcal{V}\,. \tag{15}$$

The key implication of the symmetry (14) is that then $x = H\overline{x}$ and (8) imply the following conservation law:

$$\sum_{j\in\mathcal{V}} f_{ij} = 0\,, \qquad \forall i \in \mathcal{R}\,. \tag{16}$$

Indeed, in this case, one can give the following electrical circuit interpretation: the link weights W_{ij} represent conductances and their inverses are link resistances; the value x_i is the voltage in node i; and f_{ij} is the electrical current flowing from node i to node j. Then, (15) and (16) can be read as the Ohm law and, respectively, the Kirchoff law.

Such interpretation in terms of electrical circuits has deep implications that we outline below. First, by simply verifying first-order conditions, one can show that the symmetry (14) implies that, for $k = 1, \dots, s$, the k-th column of H coincides with the solution of the following quadratic optimization problem

$$\frac{1}{R_{\mathcal{S}_k \leftrightarrow S_{-k}}} = \min_{\substack{y \in \mathbb{R}^n : \\ y_i = 1 \quad i \in \mathcal{S}_k \\ y_i = 0 \quad i \in \mathcal{S}_{-k}}} \frac{1}{2} \sum_{i,j \in \mathcal{V}} W_{ij}(y_i - y_j)^2 \,, \qquad 1 \le k \le s \,. \tag{17}$$

The quantity that is to be minimized in (17) represents the energy dissipation in the network when the voltages at the nodes are y_i. Hence, $H\delta^{(k)}$ is the vector of voltages minimizing the energy dissipation under the constraints that the voltage is 1 in all nodes on $\mathcal{S}_k$ and 0 in all nodes in $\mathcal{S} \setminus \mathcal{S}_k$. The inverse of the minimal energy dissipation in (17), i.e., $R_{\mathcal{S}_k \leftrightarrow S_{-k}}$, is known as the *effective resistance* between the node sets $\mathcal{S}_k$ and $\mathcal{S}_{-k}$. A classical duality result known as Thompson principle then implies that the inverse effective resistance between $\mathcal{S}_k$ and $\mathcal{S}_{-k}$ coincides with the total electrical current flowing out of $\mathcal{S}_k$, or, equivalently, into $\mathcal{S}_{-k}$, when all the voltages in $\mathcal{S}_k$ equal 1 and all those in $\mathcal{S}_{-k}$ equal 0, i.e.,

$$\frac{1}{R_{\mathcal{S}_k \leftrightarrow S_{-k}}} = \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{S}_k} W_{ij}(1 - H_{ik}) = \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{S}_k} W_{ij} H_{ik} \,. \tag{18}$$

One key advantage of this electrical network interpretation is that the relations (17) and (18) imply simple and powerful rules to compute or estimate voltages without the need to solve directly the Laplace equation. Particularly useful are the so called parallel and series law which allow one to replace multiple parallel links between two nodes or a line of links by a single link with conductance equal to, respectively, the sum of the conductances and the reciprocal of the sum of the reciprocal of the conductances, while maintaining the same voltages in the nodes of the electrical network. Another fundamental property implied by (17) and (18) is monotonicity of the effective resistance between two subsets of nodes with respect to changes of the network: the effective resistance is never increased when new links are added, or when the conductance, i.e., the weight, of some of the existing links is increased —including when two nodes in $\mathcal{R}$ are glued together, which can be interpreted as the addition of an infinite conductance link between them.

In fact, one can express the equilibrium states x_i in terms of effective resistances explicitly through the following formula:

$$H_{ik} = \frac{1}{2} + \frac{R_{i \leftrightarrow \mathcal{S}_{-k}} - R_{i \leftrightarrow \mathcal{S}_k}}{2R_{\mathcal{S}_k \leftrightarrow \mathcal{S}_{-k}}}$$

so that

$$x_i = \frac{1}{2} \sum_{1 \le k \le s} \overline{x}_k + \sum_{1 \le k \le s} \overline{x}_k \frac{R_{i \leftrightarrow \mathcal{S}_{-k}} - R_{i \leftrightarrow \mathcal{S}_k}}{2R_{\mathcal{S}_k \leftrightarrow \mathcal{S}_{-k}}} \tag{19}$$

Particularly simple is the case when there are just two sink components $\mathcal{S}^+$ and $\mathcal{S}^-$ with corresponding values $\overline{x}^+ = 1$ and $\overline{x}^- = 0$. In this case we obtain

$$x_i = \frac{1}{2} + \frac{R_{i \leftrightarrow \mathcal{S}^-} - R_{i \leftrightarrow \mathcal{S}^+}}{2R_{\mathcal{S}^+ \leftrightarrow \mathcal{S}^-}}$$

The sign of the difference between the two effective resistances determines if node i will be more influenced by $\mathcal{S}^+$ or $\mathcal{S}^-$. In this sense, formula (19) expresses the bias of a node towards a sink component as determined by the electrical resistance to that sink in comparison to the others. Such electrical equivalence is a classical result that has found several applications in the context of Markov chains [2, 5] and, more recently, for designing efficient distributed algorithms for the optimal stubborn node placement problem.

4 Polarization and homogeneous influence in networks

In this section, we consider graphs with $s \geq 2$ sink components all of which are singletons, $\mathcal{S}_1 = \{v_1\}, \dots, \mathcal{S}_s = \{v_s\}$, and hold different values $\overline{x}_1, \dots, \overline{x}_s \in [0, 1]$. We will investigate conditions —on the graph structure and on the size of the sink components— under which most of the entries x_i of the equilibrium vector $x = H\overline{x}$ are close to a common value $\tilde{x}$ which is a convex combination of the $\overline{x}_j$s, or rather they are all close to one of the extreme values $\overline{x}_1, \dots, \overline{x}_s$. In order to formalize these notions, we will consider infinite sequences of graphs (typically of increasing size), and briefly refer to them as (large-scale) networks. Following [1], we will say that the sink components $\mathcal{S} = \{v_1, \dots, v_k\}$ have *homogeneous influence* on (the rest of) the network if,

$$\lim_{n \to +\infty} \inf_{\tilde{x}} \frac{1}{n} \left| \{ i \in \mathcal{V} : \, |x_i - \tilde{x}| < \varepsilon \} \right| = 1 \,, \qquad \forall \varepsilon > 0 \,. \tag{20}$$

On the other hand, we will refer to a network as *polarized* if

$$\lim_{n \to +\infty} \frac{1}{n} \left| \left\{ i \in \mathcal{V} : \, \min_{1 \leq j \leq s} |x_i - \overline{x}_j| < \varepsilon \right\} \right| = 1 \,, \qquad \forall \varepsilon > 0 \,. \tag{21}$$

For the sake of simplicity, as we did in the previous section, we will confine our discussion to the special case when the restriction of the graph $\mathcal{G}$ to $\mathcal{R} = \mathcal{V} \setminus \mathcal{S}$ is connected and undirected, i.e., when (14) holds true. We start with a result on an even more special graph structure $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ whose node set $\mathcal{V}$ consists of only two stubborn nodes, v_0 and v_1, and two disjoint sets of regular nodes, $\mathcal{U}_0$ and $\mathcal{U}_1$, such that: the nodes in $\mathcal{U}_0$ (respectively, $\mathcal{U}_1$) are all connected by a weight-α directed link to v_0 (v_1); the subnetwork obtained by removing the stubborn nodes from $\mathcal{G}$ is undirected and connected; the aggregate weight $\sum_{j \in \mathcal{U}_1} W_{ij}$ (respectively, $\sum_{j \in \mathcal{U}_0} W_{ij}$) of links connecting a node $i \in \mathcal{U}_0$ ($i \in \mathcal{U}_1$) to nodes in $\mathcal{U}_1$ ($\mathcal{U}_0$) is a positive constant β_0 (β_1) independent of i. In other words, we consider a network whose weight matrix W has the structure

$$W = \begin{bmatrix} \alpha & 0 \dots & \dots 0 & 0 \\ \vdots & & & \vdots \\ \alpha & A & B & 0 \\ 0 & & & \alpha \\ \vdots & C & D & \vdots \\ 0 & 0 \dots & \dots 0 & \alpha \end{bmatrix}, \qquad \begin{array}{ll} A = A' \,, & B = C' \,, \\ C = B' \,, & D = D' \,, \\ B\mathbb{1} = \beta_0 \mathbb{1} \,, & C\mathbb{1} = \beta_1 \mathbb{1} \,. \end{array} \tag{22}$$

Proposition 3. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ be a graph with weight matrix W as in (22). Assume that the two stubborn nodes are assigned the values $\overline{x}_{v_0} = 0$ and $\overline{x}_{v_1} = 1$. Let $x = H\overline{x}$*

be the associated equilibrium vector and $y_0 := \frac{1}{n_0}\sum_{i\in\mathcal{U}_0} x_i$ *and* $y_1 := \frac{1}{n_1}\sum_{i\in\mathcal{U}_1} x_i$ *be the average states in the two subsets of nodes. Then,*

$$y_0 \le \frac{1}{1+\dfrac{n_0}{n_1}+\dfrac{\alpha}{\beta_0}}\,, \qquad 1-y_1 \le \frac{1}{1+\dfrac{n_1}{n_0}+\dfrac{\alpha}{\beta_1}}\,, \qquad y_1-y_0 \le \frac{1}{1+\dfrac{\beta_0}{\alpha}+\dfrac{\beta_1}{\alpha}}\,. \tag{23}$$

Proof Since the restriction of the graph to $\mathcal{R}$ is undirected, we can use the electrical circuit interpretation of Section 3. In particular, (18) implies that

$$\frac{1}{R_{v_0\leftrightarrow v_1}} \;=\; \underbrace{\alpha\sum_{i\in\mathcal{U}_0} x_i}_{\substack{\text{current}\\ \text{in } v_0}} \;=\; \underbrace{\sum_{i\in\mathcal{U}_0}\sum_{j\in\mathcal{U}_1} W_{ij}(x_j-x_1)}_{\substack{\text{total current}\\ \text{from } \mathcal{U}_1 \text{ to } \mathcal{U}_0}} \;=\; \underbrace{\alpha\sum_{j\in\mathcal{U}_1}(1-x_j)}_{\substack{\text{current}\\ \text{from } v_1}}\,. \tag{24}$$

On the other hand, we can get a lower bound on the effective resistance by merging all nodes in $\mathcal{U}_0$ into a single node u_0, all nodes in $\mathcal{U}_1$ into a single node u_1, and applying the parallel and series law to the resulting network, getting

$$R_{v_0\leftrightarrow v_1} \ge \frac{1}{\alpha n_0}+\frac{1}{n_0\beta_0}+\frac{1}{\alpha n_1} = \frac{1}{\alpha n_0}+\frac{1}{n_1\beta_1}+\frac{1}{\alpha n_1}\,. \tag{25}$$

Then, (23) follows by substituting the identity (24) in the lefthand side of (25). ■

It follows from Proposition 3 that, for a network with the structure as in (22),

(i) if $\alpha \gg \max\{\beta_0,\beta_1\}$, then $y_0 \to 0$ and $y_1 \to 1$;

(ii) if $\alpha \ll \max\{\beta_0,\beta_1\}$, then $y_1 - y_0 \to 0$.

Observe that, since, for all $\varepsilon > 0$,

$$\frac{1}{n}|\{i\in\mathcal{U}_0:\, x_i \ge \varepsilon\}| \le \frac{y_0}{\varepsilon}\,, \qquad \frac{1}{n}|\{i\in\mathcal{U}_1:\, x_i \le 1-\varepsilon\}| \le \frac{1-y_1}{\varepsilon}\,,$$

point (i) above implies that the network is polarized if $\alpha \gg \max\{\beta_0,\beta_1\}$. The intuition behind this result is that, if $\alpha \gg \max\{\beta_0,\beta_1\}$, then, in both $\mathcal{U}_0$ and $\mathcal{U}_1$, the total weight of links towards the stubborn node v_0 (respectively, v_1) is much larger than the total weight of links to the other set of nodes. Then, for $h = 0, 1$, nodes in $\mathcal{U}_h$ get influenced by the stubborn node v_h much more than by the other set of nodes $\mathcal{U}_{1-h}$.

On the other hand, observe that $y_1 - y_0 \to 0$ does not imply homogeneous influence (it is not hard to provide counterexamples). Sufficient conditions for homogeneous influence have been proved in [1] based on finer properties of the graph $\mathcal{G}$, in particular on its mixing time, as per the following result.

Theorem 4 ([1])**.** *Let* $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$ *be a graph with sink components* $\mathcal{S}_1, \ldots, S_s$*. Let* $\mathcal{S} = \bigcup_{1\le j\le s}\mathcal{S}_j$*,* $\mathcal{R} = \mathcal{V}\setminus\mathcal{S}$*. Let* $\tilde{\mathcal{G}} = (\mathcal{V}, \tilde{\mathcal{E}}, \tilde{W})$ *be the undirected connected graph obtained from* $\mathcal{G}$ *by making all the directed links from some node in* $\mathcal{R}$ *to some node in* $\mathcal{S}$ *bidirectional and letting the modified weight matrix* $\tilde{W}$ *coincide with* W *in its* $\mathcal{R}\times\mathcal{V}$ *block, and be such that* $\tilde{W}_{ij} = W_{ji}$ *for all* $i \in \mathcal{S}$ *and* $j \in \mathcal{R}$*, and* $\tilde{W}_{ij} = 0$ *for all* $i, j \in \mathcal{S}$*. Let* $\tilde{P}$ *be the*

stochastic matrix associated to $\tilde{\mathcal{G}}$, $\tilde{\pi} = \tilde{P}'\tilde{\pi}$ *be its invariant probability vector, and* $\tilde{\tau}$ *be the mixing time of* $\frac{1}{2}(I+P)$. *Then, the stationary vector* x *satisfies*

$$\frac{1}{n}\left|\{i \in \mathcal{V} : |x_i - \tilde{x}| \geq \varepsilon\}\right| \leq \frac{\Delta}{\varepsilon n \tilde{\pi}_*}\psi(\tilde{\tau}\cdot\tilde{\pi}_{\mathcal{S}})\,, \qquad \tilde{x} = \tilde{\pi}'x\,,$$

for every $\varepsilon > 0$, *where* $\tilde{\pi}_* := \min_{i\in\mathcal{V}}\tilde{\pi}_i$, $\psi(y) := y\log(e^2/y)$, *and* $\Delta = \max_{1\leq i,j\leq s}\{\overline{x}_i - \overline{x}_j\}$.

Theorem 4 implies that influence is homogenous in networks such that the product $\tilde{\tau}\cdot\tilde{\pi}_{\mathcal{S}}$ of the mixing time and the aggregate centrality of the set of stubborn nodes is vanishing (and $n\tilde{\pi}_*$ is bounded away from 0). Such networks such that $\tilde{\tau}\cdot\tilde{\pi}_{\mathcal{S}} \to 0$ have been referred to as *highly fluid*. Examples of highly fluid networks include d-dimensional toroidal grids with $d \geq 3$ when $|\mathcal{S}| \ll n^{1-2/d}$, and expansive networks such as the Erdos-Renyi graph provided that $|\mathcal{S}| \ll n/\log n$.

We conclude this section with an application of Proposition 3 and Theorem 4, highlighting a threshold phenomenon for the global behavior of the equilibrium state vector x, with a transition from polarization to homogeneous influence at the change of a parameter. Consider two independent and identically distributed Erdos-Renyi random graphs $\mathcal{G}_0 = (\mathcal{U}_0, \mathcal{E}_0)$ and $\mathcal{G}_1 = (\mathcal{U}_1, \mathcal{E}_1)$ with parameters m and $p = \gamma m^{-1}\log m$, where $\gamma > 1$ is a constant independent of m. Hence, both $|\mathcal{U}_0| = m$ and $|\mathcal{U}_1| = m$, and distinct pairs of nodes $\{i,j\} \subseteq \mathcal{U}_h$, $h = 0, 1$, are connected by a weight-1 undirected link independently with probability p. The choice of the scaling $p = \gamma m^{-1}\log m$ with $\gamma > 1$ guarantees that both $\mathcal{G}_0$ and $\mathcal{G}_1$ are connected with high probability as m grows large. [3] Then, let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, where $\mathcal{V} = \{v_0\}\cup\mathcal{U}_0\cup\mathcal{U}_1\cup\{w_1\}$ be the graph obtained by interconnecting $\mathcal{G}_0$ and $\mathcal{G}_1$ by an arbitrary matching of $\mathcal{U}_0$ and $\mathcal{U}_1$ of weight-β links (i.e., every node in $\mathcal{U}_h$ is connected to exactly one node in $\mathcal{U}_{1-h}$ by an undirected weight-β link) and adding a directed weight-α link from each node in $\mathcal{U}_0$ to v_0 and from each node in $\mathcal{U}_1$ to v_1. Then, Proposition 3 and Theorem 4 imply that

(i) if $\alpha \gg \beta$, then the network is polarized;

(ii) if $\alpha \ll \beta \ll 1$, then influence is homogeneous.

In fact, point (i) above is implied directly from Proposition 3. On the other hand, to verify point (ii), we first recall a result in [3, Ch. 6] stating that the Erdos-Renyi random graph in the connected regime has mixing time of order $\log m$. Then, for $h = 0, 1$, consider the network $\hat{\mathcal{G}}_h$ with regular nodes $\mathcal{U}_h$ and stubborn nodes $\mathcal{S}_h := \{w_h\}\cup\mathcal{U}_{1-h}$ obtained by removing from $\mathcal{G}$ node w_{1-h} along with all the internal links of $\mathcal{G}_{1-h}$. Observe that the aggregate centrality of $\mathcal{S}_h$ in $\hat{\mathcal{G}}_h$ is equal to $(\alpha+\beta)m/(2(\alpha+\beta)m+2L)$ where L is the total number of undirected links in the Erdos-Renyi graph $\mathcal{G}_h$, that is of order $m^2p = \gamma m\log m$ with high probability. Hence, with high probability as m grows large, the network $\hat{\mathcal{G}}_h$, $h = 0, 1$, is highly fluid if $\alpha \ll \beta \ll 1$. Hence, Theorem 4 implies that influence in both $\hat{\mathcal{G}}_0$ and $\hat{\mathcal{G}}_1$ is homogeneous (in particular, when considering as boundary conditions on $\mathcal{U}_0$, and respectively $\mathcal{U}_1$, the actual equilibrium state values), so that all but a vanishing fraction of nodes in $\mathcal{U}_0$ and $\mathcal{U}_1$ have equilibrium state close to y_0 and y_1, respectively. On the other hand, Proposition 3 implies that, if $\alpha \ll \beta$, then $y_1 - y_0 \to 0$, so that influence is indeed homogeneous across the wholes network $\mathcal{G}$.

Acknowledgements

The first author is a member of the LCCC Linnaeus Center and the ELLIIT Excellence Center at Lund University. His research has been partially supported by the Swedish Research Council (VR) through the the junior research grant Information Dynamics in Large-Scale Networks.

References

[1] D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar, *Opinion fluctuations and disagreement in social networks*, Mathematics of Operation Research **38** (2013), no. 1, 1–27.

[2] P. Doyle and J. Snell, *Random walks and electric networks*, Mathematical Association of America, 1984.

[3] R. Durrett, *Random graph dynamics*, Cambridge University Press, 2006.

[4] M. Ji, G. Ferrari-Trecate, M. Egerstedt, and A. Buffa, *Containment control in mobile networks*, IEEE Transactions on Automatic Control **53** (2008), no. 8, 1972–1975.

[5] D.A. Levin, Y. Peres, and E.L. Wilmer, *Markov chains and mixing times*, American Mathematical Society, 2010.

[6] R. Montenegro and P. Tetali, *Mathematical aspects of mixing times in Markov chains*, Foundations and Trends in Theoretical Computer Science **1** (2006), no. 3, 237–354.

[7] L. Vassio, F. Fagnani, P. Frasca, and A. Ozdaglar, *Message passing optimization of harmonic influence centrality*, IEEE Transactions on Control of Network Systems **1** (2014), no. 1, 109–120.

[8] J.C. Willems, *Models for dynamics*, Dynamics Reported **2** (1989), 171–269.

[9] ———, *Control as interconnection*, Feedback Control, Nonlinear Systems, and Complexity. Festschrift at the occasion of the 60-th birthday of George Zames, vol. 202, pp. 261–275, Springer Lecture Notes in Control and Information Sciences, 1995.