



HAL
open science

Network efficiency under nonconservative diffusion

Aitor Azemar, Ernesto Estrada

► **To cite this version:**

Aitor Azemar, Ernesto Estrada. Network efficiency under nonconservative diffusion. 2024. hal-04631930

HAL Id: hal-04631930

<https://hal.science/hal-04631930v1>

Preprint submitted on 2 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Network efficiency under nonconservative diffusion

Aitor Azemar* and Ernesto Estrada†

Abstract. The standard definition of network efficiency assumes in an ad hoc way that network navigability occurs by shortest paths only. This is obviously not the case for the many diffusive processes occurring in real-world complex systems. Here we propose from first principles a network efficiency measure that accounts for network navigability in diffusive terms. In particular, we prove that this efficiency index is based on nonconservative diffusion processes on the network, which are ubiquitous in the real-world. We then investigate analytically several properties of this efficiency index and provide computational examples of its effectivity for the analysis of complex systems. In particular, we show that the new efficiency index of a network does not necessarily change monotonically with the removal of edges, like it trivially happens with the index based on shortest paths.

Key words. Efficiency index, Information-like dynamics, Markov Processes

MSC codes. 05C12, 05C50, 05C81, 37E25, 91D30

1. Introduction. Networks [14, 26], which mathematically are graphs $G = (V, E)$, are natural representations of (complex) systems [17], where the system’s entities are represented by the vertices $v \in V$ of the network and their pairwise relations are captured by means of the edges $(v, w) \in E$ of G . Arguably such networked characteristic of (complex) systems exists to facilitate the transfer of “information” among the entities of the system [5]. Here, the term “information” is used in a very general sense to capture the idea of something, from electrons to news, which is moving across the network. Therefore, a natural question about networks is: how efficient are they? But, the core of this question is “efficient” for what? Herbert Simon, who won the Nobel prize in Economics, stressed that [33]: “In its broadest sense, to be efficient simply means to take the shortest path, the cheapest means, toward the attainment of the desired goals”. This immediately triggers another question: Shortest paths for what? In a connected network there is a topological shortest path connecting every pair of vertices. This would be the motivation of Latora and Marchiori [25] when they “introduce the concept of efficiency of a network, measuring how efficiently information is exchanged over the network”, as they defined such quantity by $E(G) = \frac{1}{n(n-1)} \sum_{v \neq w} \frac{1}{d_{v,w}}$, where $d_{v,w}$ is the shortest path distance between v and w , and n is the total number of vertices in G . We note in passing that $H(G) = \sum_{v \neq w} \frac{1}{d_{v,w}}$ was introduced in 1993 as the Harary index of a graph [31, 23]. However, the fact that every pair of vertices in G is connected by a shortest path does not mean that they are used to transmit the information between those pairs of vertices, or even that they represent the shortest way—in terms of cost, and time—to do so. It is worth reminding here some guiding words by Kurt Lewin about shortest paths [27]: “The property which makes one path between two regions of the life space outstanding

*School of Mathematics & Statistics, University of Glasgow, G12 8QQ, UK

†Institute for Cross-Disciplinary Physics and Complex Systems, Universitat de les Illes Balears, 07122, Spain

(the 'shortest path') seems to vary considerably with the situation. Sometimes the fastest connection is outstanding; at other times it is the cheapest connection' or the most pleasant, or the least dangerous. So, all that we would like to assume is that there should be, in the given case, one outstanding path from a to b." In fact, as it has been remarked by Goñi et al. [21] "a routing/navigation process implies that communication flows from a specific source to a specific target along the fastest or most direct route, which implies global knowledge about the network topology." Such global knowledge about the network topology makes the exclusive use of shortest path an impossible mission in network navigation.

Diffusion is an ubiquitous process in nature where [21] "communication occurs in the absence of specific targets, or that, even if targets are specified, a lack of knowledge about global network topology prevents particles or messages from taking shortest paths." Therefore, many dynamical processes, ranging from information propagation, synchronization and epidemic spreading, occurring in networks are diffusive in nature [29]. This diffusive processes, as commonly considered for the study of networks, is conservative in the sense that the total amount of diffusive material is constant on the graph at any time [16]. However, in many physical processes occurring in complex systems: (i) some information is dissipated to or taken from outside the graph, (ii) part of the information is annihilated/created at the vertices of the graph, which make the process nonconservative (NC) on the network [16]. Some specific examples are the traffic in a city where the number of cars flowing through the streets (edges) of an urban street network. The NC character of urban traffic emerges from the fact that some cars may "disappear"/"appear" in the street leg between two intersection because they may park or emerge from parkings in such street leg [3, 4]. Another example of NC diffusion is the flow of matter and energy in a food web. In this case the nodes represent species and the directed edges their trophic relations (who eats who), which are the pathways over which energy and matter can flow. When one species A (predator) predate another species B (prey), A utilizes only a portion of the material and free energy originally in B , which is then retained in the predator, giving rise to a non-conservative process [6]. Such mass and energy can then be diffused across the food web in an NC diffusive way. Another NC diffusion is the chemical synapses in neuronal systems where the so-called volume transmission (VT) uses the extracellular fluid filling channels of the extracellular space and the cerebrospinal fluid filling ventricular space and sub-arachnoidal space [1, 34, 35]. More examples of NC diffusion include the communication in social media like Twitter where an user can post a message who can be read by her followers, but also (if not constrained by the user) by non-followers, all of whom can retweet such information to others [28, 38, 30]. Consequently, here we ask the question: How efficient a network is in the NC diffusion of information between pairs of its vertices?

To confront the problem of quantifying the efficiency of a network in the NC diffusion of information we build on the basis of the concept of efficiency used in chemistry and biology. In a chemical reaction, the "reaction mass efficiency" [9] is defined as the percentage of actual mass of desired product to the mass of all reactants used. That is, $E_R = \frac{\text{actual mass of desired product}}{\text{mass of reactants}}$, which is frequently expressed as percentage. On the other hand, theoretical biologist Michael E. Conrad in his highly influential book "Adaptabil-

ity” defined organism efficiency as [8]: $E_O = \frac{Y}{I}$, where Y is the yield, and I is the food or light energy absorbed by the organism. Here, we propose to quantify the efficiency of a network by considering the NC diffusion of information from one vertex $v \in V$ to another vertex $w \in V$. If we allocate the total amount of information at an initial time $t = 0$ at vertex v , the “actual mass of desired product” or “yield” is the amount of information transferred to vertex w at a time $t \rightarrow \infty$, i.e., the difference between the amount of information at vertices v and w . As the process is NC, not all the information allocated at the vertex v is transferred to the rest of vertices of G . Then, the equivalent to the “mass of reactants” or “food or light energy absorbed” is given by the sum of the amounts of information at both vertices when $t \rightarrow \infty$. The efficiency measure defined and studied here is therefore different conceptually from the one proposed by Latora and Marchiori [25] (see also [36]) in the fact that it is based on a dynamical process taking place on the network without any preliminary assumption about the paths that information uses to go from one vertex to another. It also differentiates from the measures proposed by Goñi et al. [21] in which those are either based on conservative processes or on topological shortest paths. However, it is worth mentioning that these authors have given an important step forward in the definition of some efficiency measures which does take into account all the “routes” connecting pairs of vertices in a network instead of specific paths. Also, two out of three of the efficiency measures can be seen as $\left(\frac{\text{Desired output}}{\text{Total output or input}}\right)$, which is also in the spirit of the current work.

2. Preliminaries.

2.1. Network theory and linear algebra. In this section we fix the notation used through the paper, as well as give a small review of some of the notions we shall use. For an introduction to network theory and linear algebra we recommend [19]. Through this paper, G shall indicate a graph, with vertices V and edges E . We shall always let the set of vertices be $\{1, 2, \dots, n\}$. Unless mentioned, we shall assume that G is weighted with weights no larger than 1, and symmetric. A graph G is *connected* if for any two different nodes there is a sequence of adjacent edges connecting the two nodes.

We shall denote by A the adjacency matrix of G , defined such that the (i, j) -th component is equal to the weight of the connection between the node i and the node j . Note that since the graph is symmetric, so is the matrix A , and hence all the eigenvalues of A are real and A is diagonalizable. Let $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ be the eigenvalues. Then, we can write

$$A = U^T \Lambda U,$$

where U is an orthogonal matrix whose rows are the eigenvectors of A , and Λ is a diagonal matrix such that the values on its diagonal are λ_i .

For a square matrix M , the matrix exponential is defined as

$$\exp(M) = \sum_{k \geq 0} \frac{M^k}{k!}.$$

Using the diagonalization of A we have

$$\exp(A) = U^T \exp(\Lambda) U,$$

with $\exp(A)$ being a diagonal matrix with entries e^{λ_i} . Note that since all the eigenvalues of A are real and finite, all the eigenvalues of $\exp(A)$ are positive.

The i, j -th component of the matrix A^k is equal to the number of walks of length k between the nodes i and j . Hence, the components of the exponential matrix are weighted sums of the number of walks between two nodes, where larger walks are given less weight.

2.2. Pure jump Markov processes. In this section we shall recall some basic definitions related to pure jump Markov processes, needed for developing the discrete model in [section 3](#). The rest of the paper does not use the more technical results derived in that section, and does not require knowledge of Markov jump processes. As such, the reader mainly interested in the definition of the efficiency index can safely skip this subsection, as well as [section 3](#) on a first reading. For a general introduction to Markov processes the reader can check [\[12\]](#).

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $M : (\Omega, \mathbb{R}_+) \rightarrow S$ be a stochastic process adapted to the filtration \mathcal{F}_t . We shall assume that the set of states S is countable. In this setting, we say that M_t is a *Markov* process if, for any $v \in S$, $s < t$ we have $\mathbb{P}[M_t = v | \mathcal{F}_s] = \mathbb{P}[M_t = v | M_s]$. That is, if the conditioning to all the past information from the previous past point s is equivalent to conditioning to the state of the process at time s . We say that a Markov process is a *pure jump* process if for all $t \in \mathbb{R}_+$ there is almost surely some $\varepsilon > 0$ such that $M_t = M_s$ for all $s \in [t, t + \varepsilon)$. That is, if the only changes in the value of M_t are given by jumps. Finally, we say that a Markov process is *time homogeneous* if for all $s < t$, $\Delta > 0$ and $v_s, v_t \in S$ with $\mathbb{P}[M_s = v_s] > 0$ we have $\mathbb{P}[M_t = v_t | M_s = v_s] = \mathbb{P}[M_{t+\Delta} = v_t | M_{s+\Delta} = v_s]$. Note that some of the definitions for the case where S is not countable, or the process is not time homogeneous, are more technical, and again we refer to [\[12\]](#) if these are needed.

Given $s \in \mathbb{R}_+$ the stopping time $t_s(v) = \inf\{t \geq s | M_t \neq M_s\}$. Then, for $v \in S$ we have $t_s - s | M_s = v \sim \text{Exp}[\lambda(v)]$, where $\text{Exp}[\lambda(v)]$ is an exponential random variable with parameter $\lambda(v) \geq 0$. Note that by the Markov property and the time homogeneity, $\lambda(v)$ does not depend on s . We shall assume that $\lambda(v) < \infty$ for all $v \in S$, which is equivalent to requiring that no instantaneous jumps occur. We might have $\lambda(v) = 0$, in which case we have that v is an attracting point, in which the process stops evolving. Then, for any $\tilde{v} \neq v$, the transition properties are defined by $\pi(v, \tilde{v}) = \mathbb{P}[M_{t_s(v)} = \tilde{v} | M_s = v]$. Note that, again due to the Markov property and the time homogeneity, π does not depend on s . One defines then the *transition-rate matrix* by setting $q(v, \tilde{v}) = \pi(v, \tilde{v})\lambda(v)$ for $\tilde{v} \neq v$ and $q(v, v) = \pi(v, v) = -\lambda(v)$. We have

$$(2.1) \quad \mathbb{P}[M_{t+h} = \tilde{v} | M_t = v] = q(v, \tilde{v})h + o(h)$$

for all $t \geq 0$ and $\tilde{v}, v \in S$ with $\mathbb{P}[M_t = v] > 0$.

It is also possible to go the other way. That is, given a transition-rate matrix such that $\pi(v, \tilde{v}) \geq 0$ for $\tilde{v} \neq v$, and with $\pi(v, v) = -\sum_{\tilde{v} \neq v} \pi(v, \tilde{v}) > -\infty$, one can build a pure jump Markov process. However, in some cases there might be an infinite amount of jumps in a finite amount of time with positive probability, and the space of states might need to be expanded accordingly, to allow one extra point where such event occurs. If such an adjustment is needed, the resulting Markov process is called *explosive*.

Chebbi-Toumi prove in [\[7\]](#) the following condition for non-explosiveness for Markov processes in \mathbb{R}_+^k , but the same proof yields the analogous result for processes in \mathbb{N}^k . Furthermore,

note that while the paper states more hypotheses, the proof of these results does not use all of them, and just uses the hypotheses listed here.

Theorem 2.1 ([7, Lemma 2.2 and Corollary 2.3]). *Let X_t be a Markov jump process with sates in \mathbb{N}^k with transition matrix given by $q(N, S)$. Then, if*

1. $\lambda(N) \leq C_\lambda(1 + |N|), \forall N \in \mathbb{N}^k,$
2. $\max(|\tilde{N} - N| : \pi(N, \tilde{N}) > 0) \leq C, \forall N \in \mathbb{N}^k$

and X_0 is bounded almost surely, the process X_t is non-explosive and

$$\mathbb{E}[\sup_{s \leq t} |X_s|^p] < \infty$$

for all $t \geq 0$ and $p \geq 1$.

Furthermore, Chebbi–Toumi [7] prove the following equation for the expected value.

Proposition 2.2 ([7, Equation 22]). *Let X_t be a Markov jump process with sates in \mathbb{N}^k with transition matrix given by $q(N, S)$. Furthermore, let $\phi : \mathbb{N}^k \rightarrow \mathbb{N}^{k'}$ be a function growing at most polynomially. Then, if the conditions from [Theorem 2.1](#) are satisfied and X_0 is bounded,*

$$\mathbb{E}[\phi(X_t)] = \mathbb{E}[\phi(X_0)] + \int_0^t \mathbb{E}[\mathcal{L}X_s],$$

where

$$\mathcal{L}X_t = \sum_{s \in S} q(X_t, S)(S - X_t).$$

2.3. Communicability distance. In [20], Lerman and Gosh model the spread of information as a continuous phenomena resulting, under some simplifications, in the equation

$$(2.2) \quad \dot{u}(t) = -\gamma \mathcal{L}_\chi u(t),$$

where $u(t)$ is a vector in $\mathbb{R}_+^{|V|}$ representing the amount of information on each node, γ is the diffusivity coefficient and $\mathcal{L}_\chi := \chi I - A$ is the Lerman-Ghosh Laplacian with $\chi \geq 0$.

The solution to the Cauchy problem (2.2) is given by

$$(2.3) \quad u(t) = e^{-t\gamma \mathcal{L}_\chi} u^0 = e^{-t\gamma \chi} e^{t\gamma A} u^0.$$

Here we focus on the part of the solution which depends on the structure of the graph, namely $\tilde{u}(t) = e^{t\gamma A} u^0$ and hereafter will consider $\beta = t\gamma$. Let us focus on a couple of vertices designated by v and w . Then, at a given time the concentration at these nodes are:

$$(2.4) \quad \tilde{u}_v(t) = \sum_j (\exp(\beta A))_{vj} u_j^0$$

and

$$(2.5) \quad \tilde{u}_w(t) = \sum_j (\exp(\beta A))_{wj} u_j^0.$$

Then, let

$$(2.6) \quad \mathcal{F}_{vw|u_j^0=\delta_{jv}} = \tilde{u}_v|u_j^0=\delta_{jv}(t) - \tilde{u}_w|u_j^0=\delta_{jv}(t),$$

be the diffusive flow from v to w in the graph G when the initial concentration is totally located at the vertex v , $u^0(j) = \delta_{j,v}$, where $\delta_{i,j}$ is the Kronecker delta. Let us define the same in the other direction when the initial concentration is completely located at the vertex w ,

$$(2.7) \quad \mathcal{F}_{wv|u_j^0=\delta_{jw}} = \tilde{u}_w|u_j^0=\delta_{jw}(t) - \tilde{u}_v|u_j^0=\delta_{jw}(t).$$

Now, let us define the sum of the gradients between the two nodes in both directions, which is given by:

$$(2.8) \quad \mathcal{D}_{vw}(\beta A) := \mathcal{F}_{vw|u_j^0=\delta_{jv}} + \mathcal{F}_{wv|u_j^0=\delta_{jw}} = (\exp(\beta A))_{vv} + (\exp(\beta A))_{ww} - 2(\exp(\beta A))_{vw}.$$

The following result was previously proved by Estrada in [13].

Proposition 2.3 ([13, Theorem 1]). $\mathcal{D}_{vw}(\beta A)$ is a squared Euclidean distance between the pairs of vertices v and w .

The term $\mathcal{D}_{vw}(\beta A)$ is then known as the communicability distance between the corresponding vertices.

Remark 2.4. The communicability distance between two vertices in the graph represents the “traffic” waste between them. That is, if we allocate a concentration of “items” $u_v^0 = 1$ at the vertex v and wait some time t , then we find some concentration at the vertex w . If the concentrations are approximately the same, it means that the channels of communication between the two vertices are very efficient. In this case $\mathcal{F}_{vw|u_j^0=\delta_{jv}} \ll 1$. If the efficiency in the other direction (from w to v) is equally large, we have that $\mathcal{D}_{vw}(\beta A)$ is very small. On the contrary, if the efficiency of the communication channels between the two vertices is very low, the communicability distance is very large.

More precisely, we have the following expression for the communicability distance.

Proposition 2.5. *Let*

$$x_i = \sum_{1 \leq j \leq n} e^{\lambda_j} \varphi_j,$$

where φ_i are the columns of the diagonalization matrix U such that $A = U^T \Lambda U$, where Λ is a diagonal matrix, and λ_i are the eigenvalues of A . Furthermore, let y_i be the vectors formed by the columns of the matrix $\exp \beta A / 2$,

Then,

$$\mathcal{D}_{i,j} = \|x_i - x_j\|^2 = \|y_i - y_j\|^2$$

Proof. The first equality has been proven by Estrada in [13]. The second equality is proven in a similar way than the first one. Let $M_{i,j}$ be (i, j) -th component of the matrix $\exp(\beta A)$. Then,

$$M_{i,j} = \langle y_i \cdot y_j \rangle.$$

Hence,

$$D_{i,j} = \langle y_i \cdot y_i \rangle + \langle y_j \cdot y_j \rangle - 2\langle y_i \cdot y_j \rangle = \langle y_i - y_j \cdot y_i - y_j \rangle = \|y_i - y_j\|^2. \quad \blacksquare$$

3. Stochastic model of information spread. Here we present an alternative derivation of the same equation for the spread of information developed by Lerman–Gosh in [20] and explained briefly in subsection 2.3 as the expected value in an analogous stochastic model. With this we reiterate that, while originally developed for modelling the spread of information, the model also can be used for studying discrete phenomena such as the spread of diseases of a general kind. In fact, the model is a continuous time version of the discrete time model proposed by Wang–Chakrabarti–Wang–Faloutsos in [37].

Let G be a binary graph. We model the spread of information on G as a time continuous and homogeneous jump process, with possible states $\mathbb{N}^{|V|}$. Each component of the possible states represents the number of information packages on each node. The process is defined by a transition rate matrix $q(N, \tilde{N})$, which we shall make explicit shortly. Given a unit package at a node v , for each neighbouring node $u \in N_v$ there is at each small increment of time δ , a chance $\gamma\delta$ that the information is replicated there, increasing the number of packages at the node u . Such replication does not diminish the number of information packages in the node v . Equivalently, the node u has, at that same small increment of time δ , a $\gamma\delta$ probability of replicating the package from the neighbour v . This chance is repeated for each package in v , and for each neighbour of u . This part of the process is captured by setting

$$q(N, N + e_u) = \gamma \sum_{v \in N_u} n_v = \gamma \langle e_u, AN \rangle.$$

We can allow for different spreading probabilities between nodes by considering a weighted network, where the weight of an edge represents the factor by which the spread probability is multiplied. That is, a weight of $\omega_{i,j}$ for an edge between i and j represents that the probability of an information package spreading from i to j in a small time δ is multiplied by $\omega_{i,j}$. That is, $\omega_{i,j}\gamma\delta$. In this way, we get the same formula for $q(N, N + e_u)$.

Furthermore, there is also a small chance $\gamma\chi\delta$ that each information package is forgotten by the node u , resulting in a decrease in the amount of packages on u . We shall assume that χ is uniform among the network. This second process is captured by setting

$$q(N, N - e_u) = \gamma\chi n^u = \gamma \langle e_u, \chi N, \rangle.$$

All the other values of the transition matrix are set to 0, except $q(i, i)$, which is set as $-\sum_{j \neq i} q(j, i)$. For each N , the possible jumps are contained in the set of vectors \tilde{N} such that $\tilde{N} - N = \pm e_u$ for some $u \in G$. That is, the possible jumps are finite, so $\lambda(N) < \infty$ and given a starting distribution N we can build a generated jump process $M^t(N)$. As per usual, the first property we shall find is that for any starting distribution the process is non explosive. For this, we check that the conditions of Theorem 2.1 are satisfied,

Lemma 3.1. *Let $q(N, \tilde{N})$ be defined as above, and $\lambda(N) = -\sum_{\tilde{N} \neq N} (q(N, \tilde{N}))$. Then,*

1. $\lambda(N) \leq C_\lambda(1 + |N|), \forall N \in \mathbb{N}^{|V|}$,
2. $\max(|\tilde{N} - N| : \pi(N, \tilde{N}) > 0) \leq C, \forall N \in \mathbb{N}^{|V|}$

Proof. Checking the two properties is straightforward. Indeed, we have

$$(3.1) \quad \begin{aligned} \lambda(N) &= \sum_{j \in G} (q(N, N + e_j) + q(N, N - e_j)) = \gamma \sum_{j \in G} \langle e_j, AN \rangle + \gamma \sum_{j \in G} \langle e_j, \chi I N, \rangle \\ &= \gamma \langle \mathbf{1}, \chi(I + A)N \rangle \leq \gamma(\chi|N| + |E||N|), \end{aligned}$$

where $\mathbf{1}$ is the vector where each component is 1. Hence, the first property follows for $C_\lambda = \gamma(\chi + |E|)$.

For the second property we have that each possible jump is of the form $\pm e_j$, so it is satisfied for $C = 1$. ■

Then, applying Chebbi–Toumi’s [7] results [Theorem 2.1](#) and [Proposition 2.2](#) we get the following.

Corollary 3.2. *For any starting finite distribution N_0 , the process $M^t(N_0)$ is non-explosive and its components have finite p -th moments for all $p \geq 1, t \geq 0$.*

Theorem 3.3. *The expected value of the Markov process M_N^t is given by the differential equation*

$$\frac{d\mathbb{E}[M^t(N)]}{dt} = -\gamma \mathcal{L}_x \mathbb{E}[M^t(N)].$$

Proof. Letting ϕ be the identity, by Chebbi–Toumi’s [Proposition 2.2](#) we have

$$\mathbb{E}[M^t(N)] = \mathbb{E}[M^0(N)] + \int_0^t \mathbb{E}[\mathcal{L}M^s(N)].$$

By the definition of \mathcal{L} we have

$$\begin{aligned} \mathcal{L}M^t(N) &= \sum_{i \in G} (q(M^t(N), M^t(N) + e_i) - q(M^t(N), M^t(N) - e_i))e_i \\ &= \gamma \sum_{i \in G} \langle e_i, (A - \chi I_n)M^t(N) \rangle e_i = -\gamma(\chi I_n - A)M^t(N). \end{aligned}$$

Hence,

$$\mathbb{E}[\mathcal{L}M^s(N)] = -\gamma(\chi I_n - A)\mathbb{E}[M^s(N)] = -\mathcal{L}_x \mathbb{E}[M^s(N)],$$

and so the result follows. ■

Note that the only property used is the linearity of the transition probabilities with respect to the state of the Markov process, so the same proof would apply to similarly defined processes. Also from the linearity we get the following result

Theorem 3.4. *Let $N, K \in \mathbb{N}^{|V|}$. Furthermore, let $M^t(N)$ and $M^t(K)$ be two independent Markov processes defined as above. Then, the process $M^t(N+K)$ has the same law as $M^t(N) + M^t(K)$. In particular, $M^t(N) + M^t(K)$ is a pure jump Markov process.*

Proof. Fix some N and K , and let $R^t = M^t(N) + M^t(K)$. For some fixed s , define the stopping time $t_s = \inf\{t \geq s \mid R^t \neq R^s\}$. Furthermore, let t_s^N, t_s^K be the equivalent stopping times for $M^t(N)$ and $M^t(K)$. Then since the probability of $t_s^N = t_s^K$ is 0, we have almost surely

$$t_s = \min(t_s^N, t_s^K).$$

By the construction of the pure jump Markov processes, t_s^N , conditioned to the filtration \mathcal{F}_s , is an exponential random variable with parameter $\lambda(M^s(N))$, and the analogous for t_s^K . Recall that the minimum of two independent exponential random variables is again an exponential random variable, with its parameter being the sum of the two others. Hence,

$$t_s | \mathcal{F}_s \sim \text{Exp}[\lambda(M^s(N)) + \lambda(M^s(K))].$$

As seen in the proof of [Lemma 3.1](#) we have $\lambda(V) = \gamma\langle \mathbf{1}, \chi(I + A)V \rangle$, so $\lambda(V_1) + \lambda(V_2) = \lambda(V_1 + V_2)$ and we have

$$t_s | \mathcal{F}_s \sim \text{Exp}[\lambda(M^s(N) + M^s(K))] \sim t_s | R^s.$$

We shall now observe the probability distribution of R^{t_s} conditioned to \mathcal{F}_s . For a given $V \in S$, we have

$$\begin{aligned} \mathbb{P}[R^{t_s} = V | \mathcal{F}_s] &= \mathbb{P}[R^{t_s} = V | M^s(N), M^s(K)] \\ &= \mathbb{P}[t_s^N = t_s, R^{t_s} = V | M^s(N), M^s(K)] + \mathbb{P}[t_s^K = t_s, R^{t_s} = V | M^s(N), M^s(K)], \end{aligned}$$

where the second split follows from the fact that $t_s^N \neq t_s^K$ with probability 1. Let $V_1, V_2 \in S$ be any possible values for $M^s(N)$ and $M^s(K)$ respectively. Then,

$$\begin{aligned} \mathbb{P}[t_s^N = t_s, R^{t_s} = V | M^s(N) = V_1, M^s(K) = V_2] \\ = \mathbb{P}[t_s^N = t_s, M^{t_s^N}(N) = V - V_2 | M^s(N) = V_1, M^s(K) = V_2]. \end{aligned}$$

Again, by the construction of pure jump Markov processes, conditioned to M^s , the variables $M^{t_s^N}$ and t_s^N are independent. Furthermore, the event $t_s^N = t_s$ is equivalent to $t_s^N < t_s^K$, which a straightforward computation shows that has conditional probability $\frac{\lambda(M^s(N))}{\lambda(M^s(N)) + \lambda(M^s(K))}$. Hence, the last probability is equal to

$$\begin{aligned} \mathbb{P}[t_s^N < t_s^K | M^s(N) = V_1, M^s(K) = V_2] \mathbb{P}[M^{t_s^N}(N) = V - V_2 | M^s(N) = V_1, M^s(K) = V_2] \\ = \frac{\lambda(V_1)}{\lambda(V_1) + \lambda(V_2)} \mathbb{P}[M^{t_s^N}(N) = V - V_2 | M^s(N) = V_1] \\ = \frac{\lambda(V_1)}{\lambda(V_1) + \lambda(V_2)} \frac{q(V_1, V - V_2)}{\lambda(V_1)} = \frac{q(V_1, V - V_2)}{\lambda(V_1) + \lambda(V_2)}. \end{aligned}$$

Using the same argument for the other process we get

$$\mathbb{P}[R^{t_s} = V | M^s(N) = V_1, M^s(K) = V_2] = \frac{q(V_1, V - V_2) + q(V_2, V - V_1)}{\lambda(V_1) + \lambda(V_2)}.$$

We shall now see that the previous is equal to $\frac{q(V_1+V_2,V)}{\lambda(V_1+V_2)}$. If $V - (V_1 + V_2)$ is not equal to $\pm e_j$ for some j , or 0, then both sides are 0 and equality follows. If $V - (V_1 + V_2) = e_j$ then

$$q(V_1, V - V_2) = q(V_1, V_1 + e_j) = \gamma\langle e_j, AV_1 \rangle,$$

so $q(V_1, V - V_2) + q(V_2, V - V_1) = q(V_1 + V_2, V_1 + V_2 + e_j) = q(V_1 + V_2, V)$. The analogous computation for when $V - (V_1 + V_2) = -e_j$ gives the same result, and so the general equality follows. Therefore, the process is indeed a pure jump Markov process, with the same transition probabilities than M^t , and starting value $N + K$. Hence, the result follows. \blacksquare

This can be interpreted as the spread of each package of information being independent of the spread of the others. From this, as well as the finite moments, the law of large numbers can be applied, obtaining the following.

Theorem 3.5. *Let $u(t, C)$ be the solution of equation (2.2), imposing the starting condition $u(0, C) = C$. Furthermore, let $(N^n) \subset \mathbb{N}^{|V|}$ be a sequence of starting values with $|N^n| \rightarrow \infty$ and $\lim_{n \rightarrow \infty} \frac{N^n}{n} = C$. Then, for any $t > 0$ the sequence of random variables $\overline{M}_n^t = \frac{M^t(N^n)}{|N^n|}$ converges almost surely to $u(t, C)$.*

Proof. For each $j \in G$ and $i \in \mathbb{N}$ let $M_i^t(e_j)$ be an i.i.d. copy of $M^t(e_j)$. Letting N_i^n be the i -th component of N^n we have, by Theorem 3.4, that $M^t(N^n)$ has the same distribution as

$$\sum_{j \in G} \sum_{1 \leq i \leq N_j^n} M_i^t(e_j) = \sum_{j \in G} \overline{X}_j^n.$$

That is, $M^t(N^n)$ can be written as the sum of $|V|$ many independent random variables. Each of these variables is the sum of N_j^n i.i.d. random variables, each with expected value $\mathbb{E}[M^t(e_j)]$. Hence, by the strong law of large numbers, if $N_j^n \rightarrow \infty$,

$$\frac{\overline{X}_j^n}{N_j^n} \rightarrow \mathbb{E}[M^t(e_j)]$$

almost surely. If N_j^n does not go to infinity, then $\frac{N_j^n}{n} \rightarrow 0$, so

$$\frac{\overline{X}_j^n}{n} \rightarrow 0.$$

Then, recalling that if $N_j^n = 0$ we have $\overline{X}_j^n = 0$,

$$\frac{\sum_{j \in G} \overline{X}_j^n}{n} = \sum_{j \in G | N_j^n \neq 0} \frac{N_j^n}{n} \frac{\overline{X}_j^n}{N_j^n} \rightarrow \sum_{j \in G | C_j \neq 0} C_j \mathbb{E}[M^t(e_j)] = \sum_{j \in G | C_j \neq 0} C_j e^{-t\gamma \mathcal{L}_x(e_j)} = e^{-t\gamma \mathcal{L}_x C}$$

where the convergence is almost surely. \blacksquare

Note that results similar to the last one can be obtained in some cases without using the linearity property.

Using this discrete model and a similar reasoning than the one detailed in [subsection 2.3](#), the value $\exp(\beta A)_{i,j}$ can be interpreted as the average amount of instances of information that the node j has at time t , for each original information package in the node i . Consequently $\exp(\beta A)_{i,i} - \exp(\beta A)_{j,i}$ can also be interpreted as the average balance at the node i between the number of repetitions of packages originating from the node i and the different packages originating from the node j , assuming the process started with 1 package on each node.

4. Efficiency index. Following the comment on [Remark 2.4](#) we consider the sum of all distances as a measure of the total amount of repeated information within a network. Precisely, we define the following.

Definition 4.1. Let $\mathcal{D}_{vw}(\beta A)$ be the communicability distance between the vertices v and w of G . The communicability redundancy the graph is defined as

$$(4.1) \quad \mathcal{R}(G, \beta) := \frac{1}{2} \sum_{v,w \in G} \mathcal{D}_{vw}^2(\beta A).$$

Remark 4.2. According to [Remark 2.4](#), the communicability redundancy represents a measure of the global redundancies of the graph to transmit information between every pair of vertices. A small value of $\mathcal{R}(G, \beta)$ indicates a small total amount of repeated information, while a large value of the index represents a large amount of redundancies.

To define a notion of efficiency for a graph based on the dynamics of information-like phenomena we now need a notion of total information generated. Similarly to the way the redundant information between two nodes is defined to be $\exp(\beta A)_{i,i} + \exp(\beta A)_{j,j} - 2\exp(\beta A)_{i,j}$, we define the following.

Definition 4.3. Let G be connected graph, and i, j two nodes in a graph. Then, the pairwise communication between two nodes is defined as

$$\mathcal{N}_{i,j}(\beta A) = \exp(\beta A)_{i,i} + \exp(\beta A)_{j,j} + 2\exp(\beta A)_{i,j}.$$

The same proof than in [Proposition 2.5](#) yields the following result.

Lemma 4.4. Let x_i and y_i be as in [Proposition 2.5](#). Then,

$$\mathcal{N}_{i,j}(\beta A) = \|x_i + x_j\|^2 = \|y_i + y_j\|^2.$$

Now, the *total pairwise communication* $\mathcal{C}(G, \beta)$ of a graph can be defined as the sum of all the pairwise communications. That is, $\mathcal{C}(G) = \sum_{i,j} \mathcal{N}_{i,j}$. Finally, the relative amount of redundant information can be defined as the quotient between the absolute redundancy, and the total communication, and the efficiency is then defined as 1 minus the previous amount. That is, we define the following.

Definition 4.5. Let G be a connected graph. The *efficiency* of the graph is

$$\mathcal{E}(G, \beta) = 1 - \frac{\mathcal{R}(G, \beta)}{\mathcal{C}(G, \beta)} = 1 - \frac{\sum_{i,j} \mathcal{D}_{i,j}^2(\beta A)}{\sum_{i,j} \mathcal{N}_{i,j}(\beta A)}$$

Remark 4.6. While it is more practical to omit the parameter χ from the definitions of $\mathcal{R}(G, \beta)$ and $\mathcal{C}(G, \beta)$, we could have included them, resulting in

$$\mathcal{R}(G, \beta, \chi) = e^{-\beta\chi}\mathcal{R}(G, \beta)$$

and

$$\mathcal{C}(G, \beta, \chi) = e^{-\beta\chi}\mathcal{C}(G, \beta).$$

Following this reasoning, the resulting formula for the efficiency would be

$$\mathcal{E}(G, \beta, \chi) = 1 - \frac{\mathcal{R}(G, \beta, \chi)}{\mathcal{C}(G, \beta, \chi)} = 1 - \frac{\mathcal{R}(G, \beta)}{\mathcal{C}(G, \beta)} = \mathcal{E}(G, \beta).$$

That is, the efficiency of the graph is independent of the parameter χ .

4.1. Alternative expressions and rough bounds. We shall first observe the following straightforward result.

Lemma 4.7. *Let G be a graph and $\beta \geq 0$. Then,*

$$0 < \mathcal{E}(G, \beta) < 1.$$

Proof. Directly from the definitions we have $0 < \mathcal{R}(G, \beta) < \mathcal{C}(G, \beta)$, so the result follows. ■

Let $M^\beta = \exp(\beta A)$. Then, extending the definition in Benzi and Klymko [2] to consider the parameter β we define the total communicability of a vertex as:

$$(4.2) \quad TC(i, \beta) := \left(M^\beta \mathbf{1}^T \right)_i.$$

Also, the so-called Estrada index of a graph is defined as [11, 22, 15]: $EE(G, \beta) := \text{tr}(M^\beta)$, where tr is the trace of a matrix. In [10] the following index was studied: $TC(G, \beta) := \sum_i TC(i, \beta)$ with $\beta = 1$. Notice that if we define J to be the all ones matrix, then $TC(G, \beta) := \text{tr}(JM^\beta)$.

The expressions defined in the previous subsection can be expressed in terms of these two quantities.

Lemma 4.8. *Let G be a graph and $\beta \geq 0$. Then,*

$$\mathcal{R}(G, \beta) = nEE(G, \beta) - TC(G, \beta)$$

and

$$\mathcal{C}(G, \beta) = nEE(G, \beta) + TC(G, \beta).$$

Proof. The proof is rather straightforward. For a given $\beta \geq 0$, let $M = \exp(\beta A)$ be the exponential matrix. Then,

$$\mathcal{R}(G, \beta) = \frac{1}{2} \sum_{i,j \in G} \mathcal{D}_{ij}(\beta A) = \sum_{i,j \in G} M_{i,i}^\beta + M_{j,j}^\beta - 2M_{i,j}^\beta.$$

Grouping the positive terms and the negative ones the previous expression is equal to

$$\frac{1}{2} \left(\sum_{i \in G} 2nM_{i,i}^\beta - \sum_{i,j} 2M_{i,j}^\beta \right) = nEE(G, \beta) - TC(G, \beta).$$

The same proof shows the result for $\mathcal{C}(G, \beta)$. ■

This allows us to find the following bounds.

Lemma 4.9. *Let $\mathcal{R}(G)$ and $\mathcal{E}(G)$ be the redundancy and the efficiency index as defined before, and let $\lambda_1 \leq \dots \leq \lambda_n$ be the eigenvalues of A . Then,*

$$(4.3) \quad n \sum_{j=2}^n e^{\beta\lambda_j} \leq \mathcal{R}(G) \leq n \sum_{j=1}^{n-1} e^{\beta\lambda_j}$$

and

$$(4.4) \quad \frac{2e^{\beta\lambda_1}}{EE(G, \beta) + e^{\beta\lambda_n}} \leq \mathcal{E}(G) \leq \frac{2e^{\beta\lambda_n}}{EE(G, \beta) + e^{\beta\lambda_1}}.$$

Proof. First, we express $\mathcal{R}(G, \beta)$ in terms of the $EE(G, \beta)$ and $TC(G, \beta)$ as follows

$$(4.5) \quad \mathcal{R}(G, \beta) = nEE(G, \beta) - TC(G, \beta).$$

Then, let us define J to be the all ones matrix, such that $TC(G, \beta) := \text{tr}(JM^\beta)$. Now because $\lambda_n(A) \text{tr}(B) \leq \text{tr}(AB) \leq \lambda_1(A) \text{tr}(B)$ we have that

$$(4.6) \quad nEE(G, \beta) - ne^{\beta\lambda_1} \leq \mathcal{R}(G, \beta) \leq nEE(G, \beta) - ne^{\beta\lambda_n}$$

which proves the first result.

The second result is obtained similarly. We have

$$\mathcal{E}(G, \beta) = 1 - \frac{nEE(G, \beta) - TC(G, \beta)}{nEE(G, \beta) + TC(G, \beta)} = \frac{2TC(G, \beta)}{nEE(G, \beta) + TC(G, \beta)},$$

so using the same inequalities as before the result follows. ■

4.2. Upper Bound. In this section we prove a precise general lower bound for the communicability distance, which results in an upper bound for the efficiency. Recall that we are denoting as $\mathcal{D}_{i,j}$ the communicability distance.

Lemma 4.10. *Let A be a symmetric real $n \times n$ matrix with non-negative entries, with ones on the diagonal and such that each entry outside the diagonal is at most 1. Furthermore, let $x_{i,j}^t$ be the i, j -th component of the matrix e^{At} . Then, for any $0 \leq i, j \leq n$ different we have*

$$x_{i,i}^t + x_{j,j}^t - 2x_{i,j}^t \geq 2.$$

Furthermore, we have equality if and only if the corresponding columns (or rows) of A are equal.

Proof. Let u be the n -th dimensional vector defined by $u_i = 1$, $u_j = -1$ and $u_k = 0$ otherwise. Then we define

$$f(t) = x_{i,i}^t + x_{j,j}^t - 2x_{i,j}^t = u^T e^{At} u.$$

Since $e^0 = I_n$ we have $f(0) = a_{i,i} + a_{j,j} = 2$. We have

$$f'(t) = u^T A e^{At} u.$$

Hence, for $t = 0$ we have

$$f'(0) = u^T A u = a_{i,i} + a_{j,j} - 2a_{i,j} = 2(1 - a_{i,j}) \geq 0.$$

For the second derivative we have

$$f''(t) = u^T A^2 e^{At} u = \|A e^{At/2} u\|^2 \geq 0,$$

where we have used that the matrix A is symmetric and commutes with e^{At} . Hence, since $f'(0) \geq 0$ and $f''(t) \geq 0$ for all t , we have $f(t) \geq f(0) = 2$.

If the columns are equal, then $u^T A = 0$, so $f'(t) = 0$ for all t , so $f(t) = 2$ for all t . On the other hand $f''(0) = \|Au\|^2$, so $f''(0) = 0$ if and only if $Au = 0$, that is, if the columns are the same. ■

As a corollary of the previous lemma, we get the lower bound for the communicability distance.

Theorem 4.11. *Let G be an undirected weighted graph, with weights at most $\alpha > 0$, and let $i, j \in V(G)$ be two vertices of G . Then, unless i and j have the same neighbourhood and are connected by an edge of weight α , we have*

$$\mathcal{D}_{i,j}(\beta A) > \mathcal{D}_{i,j}^S(\alpha\beta) = \sqrt{2/e^{\alpha\beta}},$$

where $\mathcal{D}_{i,j}^S(\alpha\beta)$ is the communicability distance between the nodes i, j in any unweighted graph such that they have the same neighbourhood and are connected by an edge of weight α .

Proof. Let A be the adjacency matrix of G . Then, $B = A/\alpha + I_n$ satisfies the hypotheses of [Lemma 4.10](#). Furthermore,

$$e^{\alpha\beta B} = e^{\beta A + \alpha\beta I_n} = e^{\alpha\beta} e^{\beta A}.$$

Therefore, $e^{\alpha\beta} \mathcal{D}_{i,j}^2(\beta A) = x_{i,i}^\beta + x_{j,j}^\beta - 2x_{i,j}^\beta \geq 2$, where $x_{i,j}^\beta$ are the components of the matrix $\exp(\beta A)$. Hence, the inequality follows. From the lemma we have equality if and only if the columns i and j of the matrix B are the same, or, equivalently, if the nodes i and j have the same neighbourhood and are connected with an edge of weight α . ■

In particular, for an unweighted undirected graph we have

$$\mathcal{D}_{i,j}(\beta A) \geq \sqrt{2/e^\beta}.$$

As a corollary, we get that the the minimum amount of edges needed to minimise the communicability distance between two nodes is the difference between the neighbourhoods. Another corollary is the following.

Corollary 4.12. *Let K_n be the complete graph of n nodes, and let G be a non-complete weighted graph of n nodes with maximum weight $\alpha > 0$. Then,*

$$\mathcal{R}(G, \beta) > \mathcal{R}(K_n, \alpha\beta) = \frac{n(n-1)}{e^{\alpha\beta}}$$

Proof. In the complete graph all nodes have the same neighbourhoods, so by the previous theorem we have

$$\frac{1}{2} \sum_{i,j \in G} \mathcal{D}_{i,j}^2(\beta A) \geq \frac{1}{2} \sum_{i \neq j \in K_n} (D_{i,j}^S)^2(\alpha\beta) = \binom{n}{2} 2/e = \frac{n(n-1)}{e^{\alpha\beta}}.$$

Furthermore, we have equality if and only if all nodes in G have the same neighbourhood and are connected between each other with the same weight. This is only true if G is the complete graph. ■

The same proof results in the following similar result, which was conjectured by Estrada in [13, Conjecture 1].

Corollary 4.13. *Let G be a non-complete binary graph of n nodes. Then,*

$$\sum_{i,j \in G} \mathcal{D}_{i,j} > \sum_{i \neq j \in K_n} \mathcal{D}_{i,j}^{K_n}.$$

By the equality from Lemma 4.8 we get the following result.

Corollary 4.14. *Let G be a non-complete binary graph of n nodes and $\beta \geq 0$. Then,*

$$nEE(G, \beta) - TC(G, \beta) > nEE(K_n, \beta) - TC(K_n, \beta) = \frac{n(n-1)}{e^\beta},$$

Corollary 4.12 can be interpreted as the complete graph being the network with less redundancy over the networks on n nodes. Furthermore, the total pairwise communication of the network, $\mathcal{C}(G, \beta)$ is nondecreasing on the components of the adjacency matrix, with the complete graph having the largest value. Hence, one gets the following result.

Corollary 4.15. *Let G be a non-complete graph of n nodes. Then,*

$$\mathcal{E}(G, \beta) < \mathcal{E}(K_n, \beta)$$

Proof. By Corollary 4.12 we have

$$\mathcal{E}(G, \beta) = 1 - \frac{\mathcal{R}(G, \beta)}{nEE(G, \beta) + TC(G, \beta)} < 1 - \frac{\mathcal{R}(K_n, \beta)}{nEE(G, \beta) + TC(G, \beta)}.$$

Finally, the quantity $nEE(G, \beta) + TC(G, \beta)$ is a positive weighted sum of the the elements of the matrix $\exp(\beta A)$, all of which are non-decreasing on the edges, with at least some increasing. Hence, $nEE(G, \beta) + TC(G, \beta) > nEE(K_n, \beta) + TC(K_n, \beta)$, and the result follows. ■

4.3. Dependency on β . As we shall explain more in detail in [subsection 6.1](#), the parameter β can be interpreted as the reciprocal of the level of external stress to which a network is submitted (see also [18] where β is interpreted as the inverse temperature of a thermal bath in which the network is submerged). Hence, there is interest in a study of the dependency of the network efficiency on this parameter. We shall begin by observing the limits, getting the following result.

Proposition 4.16. *Let G be connected a network. Then,*

$$\lim_{\beta \rightarrow 0} \mathcal{E}(G, \beta) = \frac{2}{n+1}$$

and

$$\lim_{\beta \rightarrow \infty} \mathcal{E}(G, \beta) = \frac{2 \sum_{i,j \in G} v_n^i v_n^j}{n + \sum_{i,j \in G} v_n^i v_n^j} \leq 1,$$

where the last inequality is an equality if and only if G is a regular graph.

Proof. As $\beta \rightarrow 0$, the matrix $\exp(\beta A)$ converges to the identity, so both $EE(G, \beta)$ and $TC(G, \beta)$ converge to n , and $\mathcal{E}(G, \beta)$ converges to $\frac{2}{n+1}$.

Let Λ be the diagonal matrix formed by the eigenvalues of the adjacency matrix, and λ_n be the largest of such eigenvalue. Since G is connected, λ_n has multiplicity 1, so we have

$$\lim_{\beta \rightarrow \infty} \frac{\exp(\beta A)}{\exp(\beta \lambda_n)} = U^T \lim_{\beta \rightarrow \infty} \exp(\beta(\Lambda - \lambda_n I_n)) U = v_n v_n^T,$$

where v_n is the normalized eigenvector with eigenvalue λ_n . Then,

$$\lim_{\beta \rightarrow \infty} \frac{nEE(G, \beta)}{\exp(\beta \lambda_n)} = n \sum_{i \in G} (v_n^i)^2 = n$$

and

$$\lim_{\beta \rightarrow \infty} \frac{nTC(G, \beta)}{\exp(\beta \lambda_n)} = \sum_{i,j \in G} v_n^i v_n^j,$$

and so the limit follows.

To see that the upper limit is only an equality for regular graphs, we write

$$\sum_{i,j \in G} v_n^i v_n^j = \langle v_n, \sum v_n^i \mathbf{1} \rangle.$$

By Cauchy–Schwartz, the previous is smaller equal than

$$\|v_n\| \left\| \sum v_n^i \mathbf{1} \right\| = \left\| \sum v_n^i \mathbf{1} \right\|^2 = \sqrt{n} \langle v_n, \mathbf{1} \rangle.$$

Applying Cauchy–Schwartz again, the previous is smaller equal than

$$\sqrt{n} \|v_n\| \|\mathbf{1}\| = n.$$

In both cases we have equality if and only if the associated vectors are positive multiples of each other. In particular, for the second case we get that equality implies that v_n is a multiple of $\mathbf{1}$. This is true only for regular graphs. Then, if v_n is a multiple of $\mathbf{1}$, the first inequality also becomes an equality, and so the result follows. ■

Both limits have clear physical interpretations. A near 0 value of β implies that the pairs of vertices are hardly communicating to each other, and so the network is not providing any work. As a result, almost all information is redundant, and so the efficiency is minimal. The upper limit corresponds to a large value of β . For such value, all communications are instantaneous, and so the particular routes provided by the network are not relevant. As such, the relevant feature becomes the amount of information repeated by each node, and so we get the maximum efficiency when each node repeats the same amount of information. That is, when the graph is regular.

We shall see now that the first derivative at 0 is, for large n , close to the edge density.

Lemma 4.17. *Let G be a graph and let E be the edges of G . Then,*

$$\left. \frac{d}{d\beta} \mathcal{E}(G, \beta) \right|_{\beta=0^+} = \frac{|E|}{(1+n)^2}.$$

Proof. We have

$$(4.7) \quad \left. \frac{d}{d\beta} \mathcal{E}(G, \beta) \right|_{\beta=0^+} = - \left. \frac{d}{d\beta} \frac{\mathcal{R}(G, \beta)}{\mathcal{C}(G, \beta)} \right|_{\beta=0^+} = - \frac{\mathcal{R}'(G, 0)\mathcal{C}(G, 0) - \mathcal{R}(G, 0)\mathcal{C}'(G, 0)}{\mathcal{C}(G, 0)^2}$$

Directly from the definition we have

$$EE(G, 0) = n,$$

$$EE(G, 0)' = \sum \lambda_n = 0.$$

and

$$TC(G, 0) = n.$$

Let $\mathbf{1}$ be the vector of ones. Then, $TC(G, \beta) = \langle \mathbf{1} \cdot \exp(\beta A) \mathbf{1} \rangle$. Hence,

$$TC(G, \beta)' = \langle \mathbf{1} \cdot A \exp(\beta A) \mathbf{1} \rangle.$$

Therefore,

$$TC(G, 0)' = \langle \mathbf{1} \cdot A \mathbf{1} \rangle = \sum_{i \in G} d_i = \frac{|E|}{2}.$$

Subbing these values in (4.7) we get the result. ■

This result can be used to refine the lower bound in [Proposition 4.16](#). Following the interpretation explained after the result, we get that, for extremely low values of β , the most important parameter used to increase the efficiency is the number of edges.

While the edge density sets the first derivative at 0, in [Figure 1](#) we compare two graphs with the same edge density that show drastically different efficiencies at different values of β . Furthermore, in [subsection 5.2](#) we go over some examples where removing edges increases the efficiency for $\beta = 1$.

Increasing the parameter β translates to lowering the stress of the system, so one would expect that the efficiency increased. We have observed this increase experimentally in all of our simulations that this is indeed the case, and hence we pose it here as a conjecture.

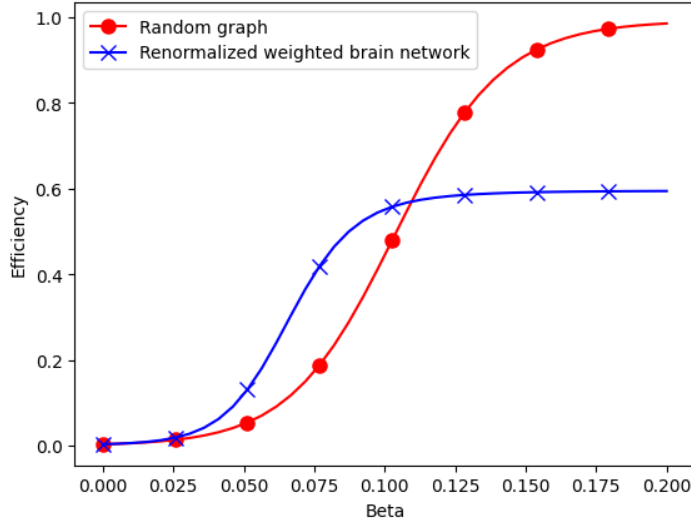


Figure 1: Efficiency value for the renormalized weighted brain coactivation network, and for an Erdős–Rényi graph with the same number of nodes and edge density than the real-world brain network.

Conjecture 4.18. *Let G be a non edgeless graph. Then, the efficiency index of G is strictly increasing on β .*

Corollary 4.19. *Let $\beta > 0$. Then, if the previous conjecture is true,*

$$\mathcal{E}(G, \beta) \geq \frac{2}{n+1},$$

with equality satisfied only for the edgeless graph of n nodes.

5. Efficiency of Networks. Artificial graphs.

5.1. Indexes comparison. In this section we investigate the general relation between the Wiener Index, the Communicability efficiency (CE) and the standard efficiency (SE) for simple graphs. In Figure 2 we have plotted the relation obtained for all connected graphs with 8 nodes. As can be seen for these small graphs, all these indices display some correlation among them. Furthermore, we observe that the relation between the WI and the CE is weaker than between the WI and the SE, which indicates that the CE is adding

While these kinds of analyses allow one to determine the relation for graphs with a small number of nodes, it is not computationally feasible to perform the same analysis for graphs with larger number of vertices. Instead, one must choose a subset of the graphs with such number of nodes. In Figure 3 we observe the relations obtained for graphs with 100 nodes, selected via the Erdős–Rényi model with $p = 0.1$, and with $p = 0.25$. We observe that the strong correlation between the Wiener Index and the SE is maintained as expected, while the relation between them and the CE weakens.

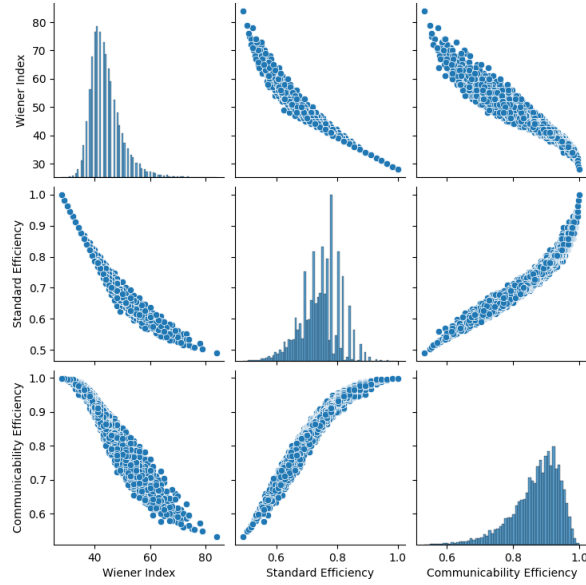


Figure 2: Relation between Wiener Index, SE and CE on graphs with 8 nodes.

The general relation between the SE and the CE indicates that these two indices observe similar phenomena. That is, while these indexes have different definitions, on many graphs they represent similar properties. However, as [Figure 3b](#) shows, and as we shall see in [section 6](#), in some families of networks the relation is broken, and the CE gives different results than the SE. In particular, we shall see that the indices behave differently on real networks associated to nonconservative dynamics. Furthermore, as observed in [Figure 3](#), there is a strong relation between WI and the SE for some families of networks, which means that the SE is not providing additional information for that family. This does not seem to be the case for the CE.

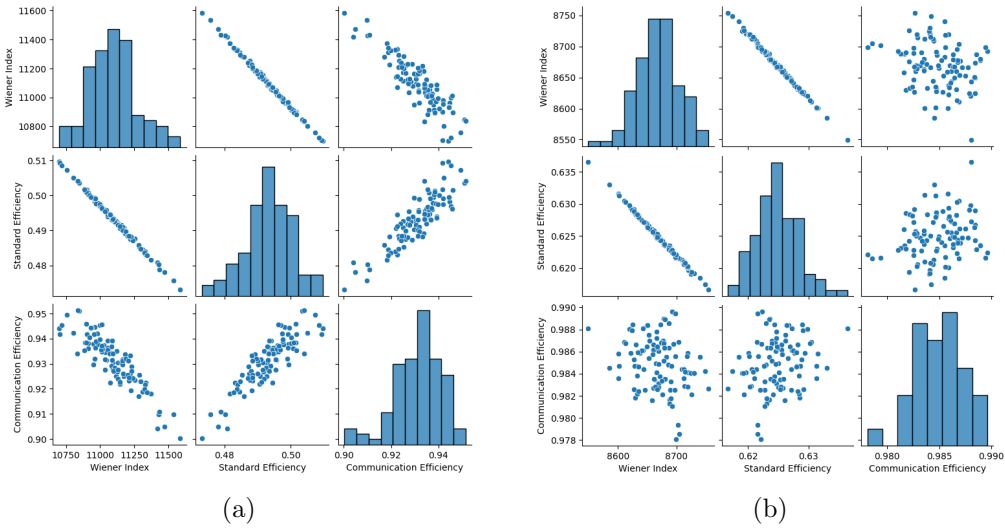


Figure 3: Relation between Wiener Index, SE and CE for 100 simulated Erdős–Rényi $n = 100$ and $p = 0.1$ (a), and for 100 simulated Erdős–Rényi graphs with $n = 100$, $k = 10$ and $p = 0.25$ (b).

5.2. Non monotony and maximally efficient subgraphs. Let G be a graph, and G' be the subgraph of G obtained by removing precisely one edge. The shortest path distance within a graph is non-decreasing with respect to edge removal. Therefore, directly from the definition one obtains that the Wiener Index and the Standard efficiency index are, respectively, increasing and decreasing with respect to edge removal. That is, $W(G) \leq W(G')$ and $\mathcal{E}^S(G) \geq \mathcal{E}^S(G')$. The same is not necessarily true for the communicability efficiency, and we shall detail an example below.

Consider the graph G formed by gluing the complete graphs of 4 and 7 vertices at a common node j . Then, a simple computation shows that removing any edge from the larger clique, except the edges adjacent to the node j , results in a graph with larger efficiency. This gives the following result.

Lemma 5.1. *The communicability efficiency is not monotonous with respect to edge removal.*

In many applications it is more cost-efficient to generate a network with a smaller number of edges, as each edge represents some form of cost. The communicability efficiency does not take into account the cost of building the network, and as such it is of interest to obtain networks with large efficiency, such that the number of edges is minimal. In particular, given a graph G , one might be interested in the subgraph of G with maximal efficiency. We define the following.

Definition 5.2. Let G be a graph. Then, we say that $H \subset G$ is a *maximally efficient*

subgraph if H contains all the nodes of G and

$$\mathcal{E}(H) = \max_{H' \subset G} \mathcal{E}(H'),$$

where the maximum is taken over the subgraphs containing all the nodes of G .

The practical applications of finding maximally efficient subgraphs leads to the following question.

Question 5.3. *Let G be a graph. Which subgraphs containing all the nodes of G are the maximally efficient subgraphs?*

We provide here the following partial result.

Lemma 5.4. *Let G be a graph with more than one connected component and let v, u be two vertices on different connected components of G . Furthermore, let G_ω be the graph obtained by adding an edge between v and u with weight $\omega > 0$. Then, for ω small enough,*

$$\mathcal{E}(G_\omega) > \mathcal{E}(G)$$

Proof. We shall first see that, for any two graphs G, \tilde{G} ,

$$\mathcal{E}(\tilde{G}) > \mathcal{E}(G) \iff \frac{TC(\tilde{G})}{EE(\tilde{G})} > \frac{TC(G)}{EE(G)}.$$

The previous fact follows from simple algebraic manipulation. Indeed,

$$\begin{aligned} & \mathcal{E}(\tilde{G}) > \mathcal{E}(G) \\ \iff & \frac{TC(\tilde{G})}{nEE(\tilde{G}) + TC(\tilde{G})} > \frac{TC(G)}{nEE(G) + TC(G)} \\ \iff & \frac{nEE(\tilde{G})}{TC(\tilde{G})} + 1 < \frac{nEE(G)}{TC(G)} + 1 \\ \iff & \frac{TC(\tilde{G})}{EE(\tilde{G})} > \frac{TC(G)}{EE(G)}. \end{aligned}$$

Consider then the function $f(\omega) = \frac{TC(G_\omega)}{EE(G_\omega)}$. Since the exponential matrix is differentiable with respect to the components of the matrix, and $EE(G) \geq 1$, the function f is differentiable. Furthermore, let $TC(\omega) = TC(G_\omega)$ and $EE(\omega) = EE(G_\omega)$. Then,

$$f'(\omega) = \frac{TC'(\omega)EE(\omega) - TC(\omega)EE'(\omega)}{EE^2(\omega)}$$

Hence, $f'(\omega) > 0$ if and only if $TC'(\omega)EE(\omega) - TC(\omega)EE'(\omega) > 0$. The value of $EE'(\omega)$ is the directional derivative of the Estrada index (see [15] for a review) of the adjacency matrix in the direction defined by the matrix with ones in the position u, v and v, u and zeros everywhere else. This quantity has been studied by Schweitzer in [32], under the name of *sensitivity* of the edge (u, v) . In particular, applying [32, Corollary 4.5] to our particular case we get

$$EE'(\omega) = 2[e^{A_\omega}]_{u,v},$$

where A_ω is the adjacency matrix of G_ω . For $\omega = 0$ we have that A_0 is a block diagonal matrix with u and v belonging to different blocks, so $e_{u,v}^{A_0} = 0$ and $EE'(0) = 0$. On the other hand, the components of e^A are nondecreasing with respect to the components of A . Hence, the derivative of TC will be larger than the derivative of the u, v component of e^{A_ω} , which observing the two first components of the definition of e^A is at least 1. Hence, $TC(0) > 0$ and $f'(0) > 0$. Therefore, $f(\omega) > f(0)$ for ω small enough, and the result follows. ■

We expect that the previous result can be extended to $\omega = 1$ for unweighted graphs, and so we pose it here as a conjecture.

Conjecture 5.5. *Let G be an unweighted graph. Then, [Lemma 5.4](#) is always true for $\omega = 1$.*

Proposition 5.6. *If [Conjecture 5.5](#) is true then, given a graph G , the connected components of the maximal efficient subgraphs of G are subgraphs containing the same vertices as the connected components of G . That is, separating two portions of a graph does not result in an efficiency increase.*

In particular, if [Conjecture 5.5](#) is true, the only maximally efficient subgraph of a forest T is T itself.

Proof. Let H be a maximally efficient subgraph of G , and let H_c be a connected component of H . Since H is a subgraph of G , the connected component H_c is a subgraph of a connected component G_c . Assume G_c has more vertices than H_c and let $u \in H_c$ and $v \in G_c - H_c$ be vertices that form an edge in G_c . Then, $H \cup (u, v)$ is still a subgraph of G , and if [Conjecture 5.5](#) is true, has higher efficiency than H , giving us a contradiction. ■

5.3. Efficiency maximizing process. While finding the maximally efficient subgraph is not computationally feasible for large graphs, here we provide an efficiency maximizing process that runs on polynomial time, given that the computation of the exponential matrix is done in polynomial time, and might be useful in some cases.

Given a network, it is possible to discern the contribution to the global efficiency provided by each edge by comparing the current efficiency with the efficiency of the graph without the observed edge. Removing the edge that contributes the least, or, in some cases, that hinders the most, repeating until we reach a totally disconnected tree, and selecting the graph with maximal efficiency, one obtains an optimising process. The resulting pseudocode is outlined in [Algorithm 5.1](#).

Here we build the following example. Similarly to the previous case, we shall work with a graph obtained by contracting two densely connected graphs at one node. This time, instead of taking complete graphs, we take two Erdős–Rényi graphs with $n = 15$ and $p = 0.7$ and $p = 0.9$. The particular graph obtained by the simulation is shown in [Figure 4a](#). After computing the change in efficiency after removing each individual edge we obtain the graph illustrated in [Figure 4b](#), indicating that the edges on the more dense subgraph are hindering the efficiency, and the edges connecting with the central node are contributing the most. Removing edges as mentioned we obtain the efficiency graph shown in [Figure 5](#). We observe that a maximal efficiency plateau is reached after removing a small amount of edges, corresponding to bringing the densities of the two clusters to similar values. After all edges that hinder efficiency are removed, the global efficiency starts decreasing with increasing pace. The approximation of the maximally efficiency graph is shown in [Figure 4c](#). Note that all connections between the

Algorithm 5.1 Greedy efficiency maximising process

Input Graph G

Output Approximation of maximally efficient subgraph of G

Declare array u containing only G

Declare array v containing only the CE of G

$F \leftarrow \text{copy}(G)$

while F has edges **do**

 Declare empty array w

for $edge \in \text{edges of } F$ **do**

 Append CE of F without $edge$ on array w

end for

 Find largest value among values in w

 Remove associated edge to previously found largest value to F

 Append a copy of F to array u

 Append CE of F to array v

end while

Find index of largest value among values in v

return Associated graph in u

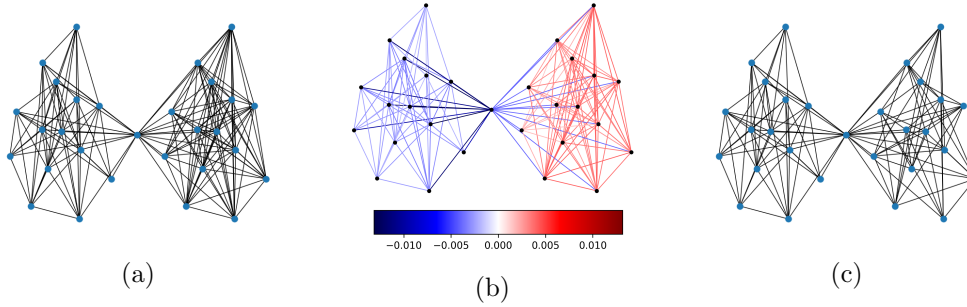


Figure 4: Example of efficiency maximizing process of edge removal sequence explained in [subsection 5.2](#). For this simulation, the maximum is attained after 42 edge removals, shown in (c). See [Figure 5](#) for the relation between edges removed and the efficiency in this example.

nodes in the sides and the central node are maintained, and that the edge density on both sides of the graph is similar, even though we have not imposed such restrictions explicitly.

We emphasise that the graph obtained through this process is not necessarily the maximally efficient graph. Furthermore, it is not enough to stop at the first local maxima, as the efficiency curve obtained as we remove the edges might have several local maxima. To illustrate this point we repeat the process described above with two complete graphs of 15 nodes, glued at a common node. Recording the efficiency of the 100 first steps of the process we obtain the curve shown in [Figure 6](#).

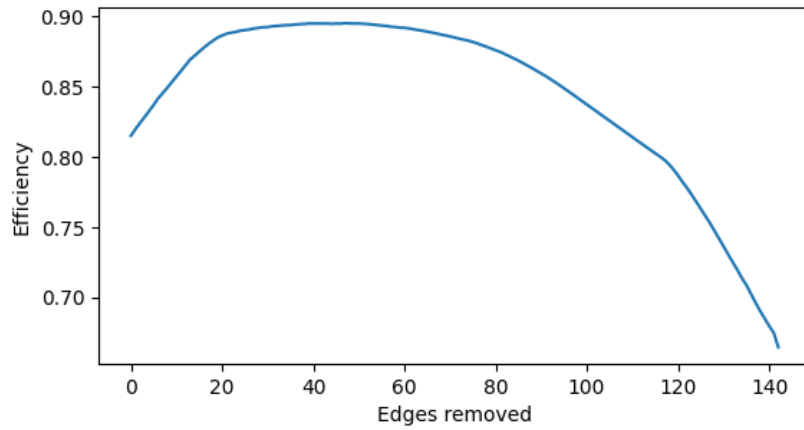


Figure 5: Efficiency obtained performing the efficiency optimisation process described in [subsection 5.2](#)

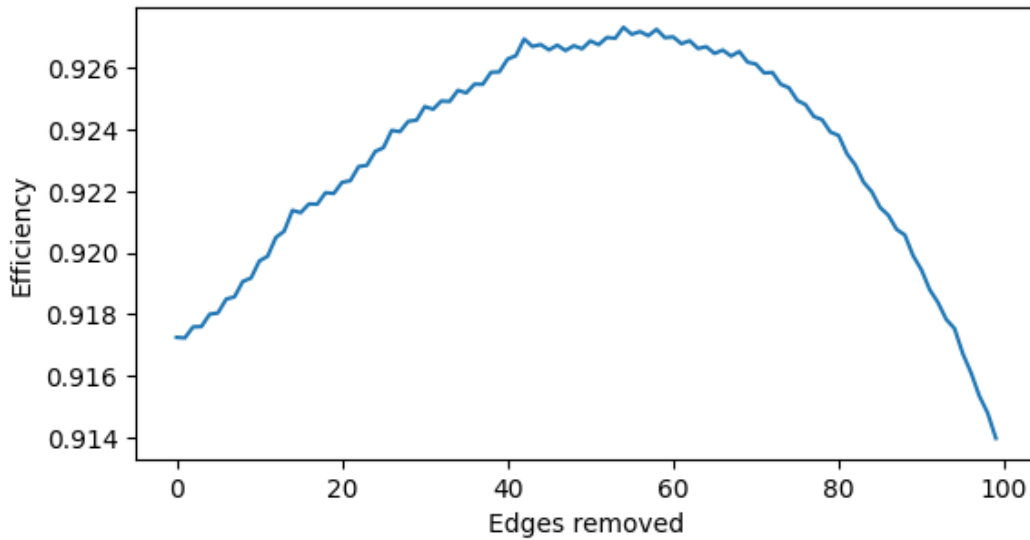


Figure 6: First 100 efficiency values obtained performing the efficiency optimisation process described in [subsection 5.2](#) for two complete graphs of 15 nodes glued at a common point.

6. Real Networks. In this section we study a series of networks from a variety of real-world complex systems. They include: (i) brain networks, (ii) food webs, (iii) online and (iv) face-to-face social networks. We selected these classes of networks because there is evidence that such systems are examples in which nonconservative diffusive dynamics take place as previously described in the Introduction. In total we consider 14 networks in these categories for which we calculated the communicability efficiency as defined in this work as well the

Network	type	Communicability Efficiency	Standard Efficiency
Human brain coactivation	i	0.6723	0.4949
Macaque Cortex	i	0.9329	0.7141
Cat cortex	i	0.9238	0.6902
C. elegans neurons	i	0.6868	0.4222
Skipwith	ii	0.9570	0.7950
St. Martin	ii	0.8856	0.5884
Bridge Brook	ii	0.8383	0.5426
Ythan	ii	0.7113	0.4574
Twitter US Congress	iii	0.8562	0.5197
Twitter Russian trolls	iii	0.3039	0.2223
social3	iv	0.7412	0.5092
Corporate People	iv	0.5722	0.3069
dolphins	iv	0.5400	0.3792
Drugs	iv	0.2849	0.2293

Table 1: Values of the communicability efficiency defined in this work as well as the standard one based on shortest path distance for a series of networks representing (i) brain networks, (ii) food webs, (iii) online and (iv) face-to-face social networks.

standard network efficiency based on the shortest path distance between pairs of vertices. The results are given in [Table 1](#). The average communicability efficiency of these networks is of about 71% in contrast with the standard efficiency which is of 49%. As average, these networks display about 21% more efficiency in the use of nonconservative diffusion than in using a hypothetical process which occurs through the shortest paths. Although both efficiency measures are based on very different principles they can be compared quantitatively because they are both magnitudes bounded between zero and one.

To put more in value the differences observed for these real-world networks in their two classes of efficiency we consider now another set of 14 networks where NC diffusion is not expected to be a major driver of their functioning. They represent systems which range from sexual intercourse networks, the galleries of termite mounds, flip-flop electronic circuits, a power-grid and several software collaboration graphs. The results are given in [Table 2](#). The average communicability efficiency of these networks is only of 16%, which is a bit smaller than the one obtained for the shortest path, which is about 19%. As average, these networks display about 3% less efficiency in the use of nonconservative diffusion than in using a hypothetical process which occurs through the shortest paths. Whether shortest paths between pairs vertices is the way in which “information” is spread through these networks could be debatable, e.g., it may happen than conservative diffusion would be a better way of doing so in certain cases. However, what it seems clear is that these networks do not work on the basis of nonconservative diffusion. For instance, a disease being transmitted through sexual contacts is not spilled over outside the intercourse interactions, neither electricity is transmitted outside the wiring connections of the circuits or the power-grid. Termite galleries seem to be prepare

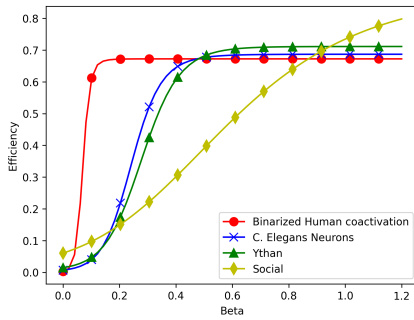
Network	Communicability Efficiency	Standard Efficiency
Heterosexual	0.2400	0.2478
Homosexual	0.0820	0.1605
termite 1	0.0460	0.1493
termite 2	0.0514	0.1451
termite 3	0.1192	0.1759
electronic1	0.1454	0.2480
electronic2	0.0815	0.2041
electronic3	0.0450	0.1693
Power Grid West USA	0.0049	0.0628
Software Abi	0.2806	0.2187
Software Digital	0.2982	0.2495
Software MySQL	0.3431	0.2265
Software VTK	0.3418	0.2445
Software XMMS	0.1707	0.1841

Table 2: Values of the communicability efficiency defined in this work as well as the standard one based on shortest path distance for a series of networks representing sexual intercourse networks, the galleries of termite mounds, flip-flop electronic circuits, a power-grid and several software collaboration graphs.

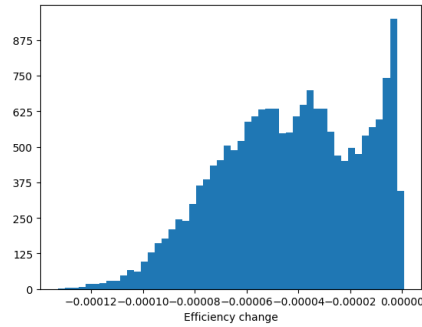
to evacuate gases possibly in a conservative diffusive way. Finally, software collaboration systems represent the way in which software modules call each other to perform tasks, which is expected to occur in an approximate shortest-path way.

In closing, the previous results may indicate that real-world networks could evolve by optimizing their structures to guarantee the efficiency of the processes to which they have evolved to perform. Therefore, networks evolved to work under nonconservative diffusive processes may have optimized their communicability efficiency in the way it is defined here. Other networks, such as the ones shown in Table 2 may have evolved by optimizing other structural features not necessarily designed for nonconservative process. The apparent ubiquity of non-conservative diffusion in complex systems make the current definition of great relevance for the analysis of realistic complex systems scenarios.

6.1. Dependency on β on real networks. We now briefly consider the effect of changing the value of β on the efficiency of a network towards NC diffusion. Let us first signify what β should mean in this context. When the matrix $\exp(\beta A)$ is obtained in reality we are computing the communicability between every pair of vertices for the matrix $W = \beta A$, which is the adjacency matrix of the graph in which every edge has the weight β . Obviously when $\beta = 1$ we have the unweighted graph, and when $\beta \rightarrow 0$ we are reducing the capacity of every edge to transmit information. For $\beta > 1$ we are increasing such capacity of communication between vertices, think for instance that if $\beta \in \mathbb{Z}$ it represents multiple connections between pairs of vertices. Therefore, as mentioned before β can be thought as the external stress to which a network is submitted to, when $\beta \rightarrow 0$ we increase the stress such that the communication



(a) Efficiency for some networks as the value of β changes



(b) Change in efficiency for the binarized reweighted brain networks as we remove edges

Figure 7

between pairs of vertices is almost interrupted.

Here we compare four real-world networks which have similar efficiency for $\beta = 1$: human brain (0.6723), *C. elegans* neurons (0.6868), the food web of Ythan (0.7113) and the social network of adolescents, social3 (0.7412). In Figure 7a we illustrate the change in the network efficiency when β drops. The first interesting observation is the fast drop of the efficiency of the social network when the external stress increases. For values of $0.2 \leq \beta \leq 0.9$ this network passes from being the most efficient to be the least one. This is intuitive on the basis of the fact that social networks are not designed to be resilient to external stress and indeed they are usually very volatile on the change of such external conditions. Think for instance how political or economical crises alterate the structure of our everyday social networks. On the other side of the coin it is the network representing the human brain coactivation system. This network seems extremely robust to the drop of β , and its efficiency only decays significantly when the external stress is dramatically large. Biological networks in general, and the brain in particular, are known to display a large adaptability and plasticity [24], allowing them to work on stressful situations. More investigations on this area are necessary to extract more generalistic conclusions but the main message of this part of the work is that “efficiency” is not only a term that needs to be used in the context of the process under consideration but also on the specific circumstances in which the complex system is operating. This concept of efficiency under external stress is different from the one typically considered in network theory where efficiency is considered on the basis of random failures of edges. We notice here (see Figure 7b) that removing edges of the human brain coactivation network namely diminish its efficiency under almost all circumstances, with only 0.8% of edge removals resulting in an efficiency increase. This indicates that the wiring is somehow necessary to keep the level of robustness and functioning of the brain.

Finally, we would like to call the attention about the following fact. In some occasions the networks under study are weighted, i.e., some real positive values are assigned to the edges with different meaning depending on the cases. Then, comparing the efficiency of a

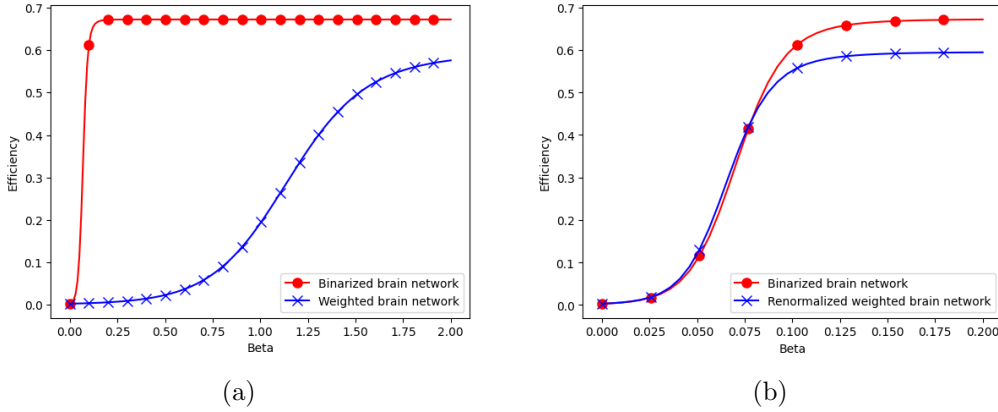


Figure 8: Comparison between the efficiencies as functions of the parameter β for the binarized brain coactivation, the weighted brain coactivation, and the renormalized brain coactivation

weighted network with the one of an unweighted one is not a straightforward task and need some thoughts. Let us consider a simple but illustrative example. A triangle in which the three edges are weighted by numbers 1, 2 and 3 have communicabilities, i.e., values of $(e^A)_{vw}$ with magnitudes of the order of 20. Comparing this network with the unweighted triangle, which has communicabilities of the order of 3 is not fair. The reason is the average weight of the edges in the first case is 2, which is similar as considering weights 0.5, 1, and 1.5 and a value of $\beta = 2$ in the first case, while in the second we are using a value of $\beta = 1$. Therefore, we should either normalize the edge weights in the weighted graph to have average equal to one in the network or find the appropriate value of β to multiply the adjacency matrix of the unweighted graph. For instance, if we consider $(e^{2A})_{vw}$ for the unweighted graph we obtain values of the order of 18.1 which are closer to the ones of the weighted graph.

The previously described situation can be seen if we compare the change of the efficiency in the human brain coactivation network with weights and the corresponding binarized network. In [Figure 8a](#) we illustrate the change of efficiency as a function of β for these two versions of the human brain network considered here. The results give the false impression of a remarkable increase of the resilience of the unweighted network relative to the weighted one. However, the appropriate normalization of the edge weights in the weighted graph makes the results comparable as seen in [Figure 8b](#) where it can be seen that there is a similar pattern in the change of the efficiency with β in both versions of the same network.

REFERENCES

- [1] Luigi F Agnati, Diego Guidolin, Michele Guescini, Susanna Genedani, and Kjell Fuxe. Understanding wiring and volume transmission. *Brain research reviews*, 64(1):137–159, 2010.
- [2] Michele Benzi and Christine Klymko. Total communicability as a centrality measure. *Journal of Complex Networks*, 1(2):124–149, 05 2013.
- [3] Ashish Bhaskar, Edward Chung, and André-Gilles Dumont. Estimation of travel time on urban networks

- with midlink sources and sinks. *Transportation research record*, 2121(1):41–54, 2009.
- [4] Ashish Bhaskar, Takahiro Tsubota, Edward Chung, et al. Urban traffic state estimation: Fusing point and zone based data. *Transportation Research Part C: Emerging Technologies*, 48:120–142, 2014.
- [5] Stefano Boccaletti, Vito Latora, Yamir Moreno, Martin Chavez, and D-U Hwang. Complex networks: Structure and dynamics. *Physics reports*, 424(4-5):175–308, 2006.
- [6] Åke Brännström and David JT Sumpter. Coupled map lattice approximations for spatially explicit individual-based models of ecology. *Bulletin of mathematical biology*, 67(4):663–682, 2005.
- [7] Mohsen Chebbi and Salwa Toumi. Law of large numbers and central limit theorem for a class of pure jump markov process. *Journal of Mathematical Analysis and Applications*, 491(1):124263, 2020.
- [8] Michael Conrad. *Adaptability: The significance of variability from molecule to ecosystem*. Springer Science & Business Media, 2012.
- [9] David JC Constable, Alan D Curzons, and Virginia L Cunningham. Metrics to ‘green’ chemistry—which are the best? *Green Chemistry*, 4(6):521–527, 2002.
- [10] Kinkar Ch. Das, Mohammad Ali Hosseinzadeh, Samaneh Hossein-Zadeh, and Ali Iranmanesh. Some bounds for total communicability of graphs. *Linear Algebra and its Applications*, 569:266–284, 2019.
- [11] Hanyuan Deng, Slavko Radenkovic, and Ivan Gutman. The estrada index. *Applications of Graph Spectra*, pages 123–140, 2009.
- [12] Andreas Eberle. Markov processes. *Lecture Notes at University of Bonn*, 2009.
- [13] Ernesto Estrada. The communicability distance in graphs. *Linear Algebra and its Applications*, 436(11):4317–4328, 2012.
- [14] Ernesto Estrada. *The structure of complex networks: theory and applications*. American Chemical Society, 2012.
- [15] Ernesto Estrada. The many facets of the estrada indices of graphs and networks. *SeMA Journal*, 79, 10 2021.
- [16] Ernesto Estrada. Conservative vs. non-conservative diusion towards a target in a networked environment. *The target problem*, 2023.
- [17] Ernesto Estrada. What is a complex system, after all? *Foundations of Science*, pages 1–28, 2023.
- [18] Ernesto Estrada and Naomichi Hatano. Statistical-mechanical approach to subgraph centrality in complex networks. *Chemical Physics Letters*, 439(1):247–251, 2007.
- [19] Ernesto Estrada and Philip A Knight. *A first course in network theory*. Oxford University Press, USA, 2015.
- [20] Rumi Ghosh and Kristina Lerman. Impact of dynamic interactions on multi-scale analysis of community structure in networks. *arXiv: Social and Information Networks*, 2012.
- [21] Joaquín Goñi, Andrea Avena-Koenigsberger, Nieves Velez de Mendizabal, Martijn P van den Heuvel, Richard F Betzel, and Olaf Sporns. Exploring the morphospace of communication efficiency in complex networks. *PLoS One*, 8(3):e58070, 2013.
- [22] Ivan Gutman, Hanyuan Deng, and Slavko Radenkovic. The estrada index: an updated survey. *Selected Topics on Applications of Graph Spectra, Math. Inst., Beograd*, pages 155–174, 2011.
- [23] Ovidiu Ivanciuc, Teodor-Silviu Balaban, and Alexandru T Balaban. Design of topological indices. part 4. reciprocal distance matrix, related local vertex invariants and topological indices. *Journal of Mathematical Chemistry*, 12(1):309–318, 1993.
- [24] Barbro B Johansson. Brain plasticity in health and disease. *The Keio journal of medicine*, 53(4):231–246, 2004.
- [25] Vito Latora and Massimo Marchiori. Efficient behavior of small-world networks. *Physical review letters*, 87(19):198701, 2001.
- [26] Vito Latora, Vincenzo Nicosia, and Giovanni Russo. *Complex networks: principles, methods and applications*. Cambridge University Press, 2017.
- [27] Kurt Lewin. Contributions to psychological theory. *The conceptual representation and measurement of psychological forces. Durham, NC, US: Duke University Press. doi*, 10:13613–000, 1938.
- [28] Lu Liu, Jie Tang, Jiawei Han, and Shiqiang Yang. Learning influence from heterogeneous social networks. *Data mining and knowledge discovery*, 25:511–544, 2012.
- [29] Naoki Masuda, Mason A Porter, and Renaud Lambiotte. Random walks and diffusion on networks. *Physics reports*, 716:1–58, 2017.
- [30] Lucas A Overbey, Benjamin Greco, Christopher Paribello, and Terresa Jackson. Structure and prominence

- in twitter networks centered on contentious politics. *Social Network Analysis and Mining*, 3:1351–1378, 2013.
- [31] Dejan Plavšić, Sonja Nikolić, Nenad Trinajstić, and Zlatko Mihalić. On the harary index for the characterization of chemical graphs. *Journal of Mathematical Chemistry*, 12:235–250, 1993.
- [32] Marcel Schweitzer. Sensitivity of matrix function based network communicability measures: Computational methods and a priori bounds. *SIAM Journal on Matrix Analysis and Applications*, 44(3):1321–1348, 2023.
- [33] Herbert Simon. *Reason in human affairs*. Stanford University Press, 1990.
- [34] E Sykova. Extrasynaptic volume transmission and diffusion parameters of the extracellular space. *Neuroscience*, 129(4):861–876, 2004.
- [35] Katherine H Taber and Robin A Hurley. Volume transmission in the brain: beyond the synapse. *The Journal of neuropsychiatry and clinical neurosciences*, 26(1):iv–4, 2014.
- [36] I Vragović, E Louis, and Albert Díaz-Guilera. Efficiency of informational transfer in regular and complex networks. *Physical Review E*, 71(3):036122, 2005.
- [37] Yang Wang, D. Chakrabarti, Chenxi Wang, and C. Faloutsos. Epidemic spreading in real networks: an eigenvalue viewpoint. In *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.*, pages 25–34, 2003.
- [38] An Zeng and Chi Ho Yeung. Predicting the future trend of popularity by network diffusion. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26(6), 2016.