

## Expected distance based on random walks

Eglantine Camby, Gilles Caporossi, Marcia Paiva, Marcelo Eduardo Vieira  
Segatto

► **To cite this version:**

Eglantine Camby, Gilles Caporossi, Marcia Paiva, Marcelo Eduardo Vieira Segatto. Expected distance based on random walks. *Journal of Mathematical Chemistry*, Springer Verlag (Germany), 2018, 56 (2), pp.618-629. hal-01944246

**HAL Id: hal-01944246**

**<https://hal.inria.fr/hal-01944246>**

Submitted on 4 Dec 2018

**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.

## Expected distance based on random walks

E. Camby · G. Caporossi · M.H.M. Paiva ·  
M.E.V. Segatto

Received: date / Accepted: date

**Abstract** By considering a graph as a network of resistances, Klein and Randić [14] proposed the definition of a distance measure. Indeed, if each edge of the graph represents a resistance of  $1\Omega$ , the equivalent resistance of the graph between each pair of vertices may be used as a distance. Based upon random walks in graphs, Stephenson and Zelen [17] built a computational model to find the probability that each edge is used. From a mathematical point of view, both articles are based upon exactly the same model and the link between random walks and the electrical representation was established by Newman [16] when defining an alternative to Freeman's betweenness centrality [9,10] based upon random walks.

In the present paper, the similitude between these two processes is exploited to propose a new random walks based distance measure that may be defined as the expected length of a walk between any pair of vertices. We call it the expected distance and we prove that it is actually a distance. From this new definition, the RW Index is proposed that sums the expected walks lengths between pairs of vertices exactly in the same way as the Wiener index sums the shortest paths distances or the Kirchhoff index sums the equivalent

---

This work was partially supported by a travel grant from the F.R.S.-FNRS, by a post-doc grant "Bourse d'Excellence WBI.WORLD" from Fédération Wallonie-Bruxelles (Belgium) and also by NSERC, Foundation HEC Montréal (Canada), CAPES, and CNPq (Brazil).

E. Camby  
Université Libre de Bruxelles, Bruxelles, Belgique  
Tel.: +1 514 340 6053 # 6032  
E-mail: [ecamby@ulb.ac.be](mailto:ecamby@ulb.ac.be)

E. Camby & G. Caporossi  
GERAD & HEC Montréal, Montréal (Québec), Canada  
E-mail: [gilles.caporossi@hec.ca](mailto:gilles.caporossi@hec.ca)

M.H.M. Paiva & M.E.V. Segatto  
Federal University of Espírito Santo, Vitória (Espírito Santo), Brazil  
E-mail: [marcia.paiva@ufes.br](mailto:marcia.paiva@ufes.br), [segatto@ele.ufes.br](mailto:segatto@ele.ufes.br)

resistances. We compare the three indices and establish the vertex and the edge decompositions for both. We compute some value of the RW index for some families of graphs and conjecture the upper and lower bounds of the RW index.

**Keywords** random walks · distance · topological index · graph

**Mathematics Subject Classification (2000)** MSC 05C81 · MSC 05C12 · MSC 05C21 · MSC 05C38

## 1 Introduction

Let  $G = (V, E)$  be a graph without loops or multiple edges on  $n = |V|$  vertices and  $m = |E|$  edges. All basic definitions can be found in [5]. A graph  $G$  is said *connected* if there exists a sequence of edges joining any pair of vertices, i.e. there exists a geodesic path between any pair of vertices. Connected graphs with minimum number of edges are called *trees*. Then, a tree on  $n$  vertices has  $n - 1$  edges. We consider in this paper only connected graphs.

The Wiener index [18], likely the oldest and most popular topological index, was originally defined for trees as:

$$\sum_{(s,t) \in E} n_s \times n_t, \quad (1)$$

where  $n_s$  is the number of vertices of the connected component containing  $s$  if the edge  $(s, t)$  was removed, and similarly for  $n_t$ . In trees, the Wiener index corresponds to the sum of distances between pairs of vertices in  $G$ :

$$W(G) = \frac{1}{2} \sum_{s \in V} \sum_{t \in V} d_{st} = \sum_{s < t \in V} d_{st}, \quad (2)$$

where  $d_{st}$  denotes the distance between  $s$  and  $t$ , i.e. the number of edges in a geodesic path joining  $s$  and  $t$ . The definition given in (2) easily extends to general graphs and is commonly used as the definition of the Wiener index.

The concept of distance is a key feature in the design of molecular descriptors. However, the distance based upon the geodesic paths has only integer values and inevitably has the same value for a large number of pairs of vertices. To overcome this weakness, Klein and Randić [14] proposed an alternative way to define the distance between pairs of vertices. Considering the graph as a network of resistors and each edge being associated to a resistor of  $1\Omega$ , they proposed to consider the equivalent resistance  $r_{st}$  between  $s$  and  $t$  as the definition of the distance between  $s$  and  $t$ . Then, analogously to the Wiener index, the sum of the equivalent resistances among all pairs of vertices corresponds to the Kirchhoff index. In practice, these equivalent resistances can be computed from the flow conservation used for electric circuits. The

latter could be described as a set of equations characterized by the Laplacian matrix  $L = D - \mathcal{A}$  where  $D$  is a diagonal matrix of degrees whereas  $\mathcal{A}$  is the adjacency matrix.

Studying random walks in social networks, Stephenson and Zelen [17] also used a matrix that turns out to be related to the Laplacian matrix. The analogy between random walks and electric circuits was later demonstrated formally by Newman [16]. Indeed, considering a unit flow of 1 ampere (1A) between two vertices  $s$  and  $t$ , the intensity on a given edge  $e$  of the network corresponds to the probability that  $e$  is used by a random walk from  $s$  to  $t$ . A general reference to random walks on graphs is given by Doyle and Snell [6].

In this paper, we introduce the expected distance between any pair of vertices as the expected length of a path between them based upon the so-defined probabilities in random walks. Instead of only considering the shortest paths between pairs of vertices as the Wiener index does, from this distance, we define a new index based on Random Walks, called RW index.

In Section 2, we present required background on random walks. In Section 3, we define the expected distance and the RW index while we compare the latter to the Kirchhoff and the Wiener indices in Section 4. In Section 5, we propound a decomposition of the RW index based upon edge or vertex centrality measures and the relation with a similar decomposition of the Wiener [4] and Kirchhoff indices. The values of the RW index for some families of graphs, conjectures on their bounds and partial results are presented in Section 6 while Section 7 concludes.

## 2 Background

As explained in the introduction, the Wiener index of a graph  $G$  is usually defined by the sum of distances between pairs of vertices:

$$W(G) = \frac{1}{2} \sum_{s \in V} \sum_{t \in V} d_{st} = \sum_{s < t \in V} d_{st} \quad (3)$$

This index gives an indication on how pairs of vertices are close. Another similar index is the Kirchhoff index. Before defining this index, we give some basic tools.

The mathematical aspects involved in the present paper are related to the use of a graph as representation of an electrical circuit. Suppose a graph  $G$  that represents a network of resistors, each edge corresponding to a unit resistor. If an electrical flow of 1A is sent through the network from vertex  $s$  to vertex  $t$ , this flow will be distributed among the circuit in a way that respects both the flow conservation equations and the electrical properties of the resistors. Indeed, according to the flow conservation rule, the sum of flows from  $s$  is 1

as well as the sum of flows to  $t$  and the net flow, i.e. the difference between inflow and outflow, associated to each other vertex is always 0. These rules are described by the following equations, where  $i_{st}^{uv}$  is the intensity of the flow from  $u$  to  $v$  when  $s$  is the source and  $t$  is the sink:

$$\sum_{k:(s,k) \in E} i_{st}^{sk} = 1 \quad (4)$$

$$\sum_{k:(t,k) \in E} i_{st}^{tk} = -1 \quad (5)$$

$$\sum_{k:(q,k) \in E} i_{st}^{qk} = 0 \quad \forall q \in \{1 \dots n\} \setminus \{s, t\} \quad (6)$$

The physical representation of the resistor is such that if an intensity  $i_{st}^{uv}$  flows through the edge  $(u, v)$  from  $u$  to  $v$ , the potentials  $d_{st}^{(u)}$  and  $d_{st}^{(v)}$  in these vertices will respect the following general equation:

$$\text{Potential Difference} = \text{Resistance} \times \text{Intensity}. \quad (7)$$

In our case, with the appropriated notation, it means that

$$d_{st}^{(u)} - d_{st}^{(v)} = i_{st}^{uv}, \quad (8)$$

since the edge  $(u, v)$  has a resistance of  $1\Omega$ . A consequence is that

$$-1 \leq d_{st}^{(u)} - d_{st}^{(v)} \leq 1 \quad \forall (u, v) \in E. \quad (9)$$

The potential  $d_{st}^{(u)}$  associated to each vertex  $u$  could be computed by the resolution of the system of equations (4)-(6) and (8). Actually, this system of equations are represented by a matrix equation, where appears the Laplacian matrix  $L$ :

$$LV = S, \quad (10)$$

where  $V = \{d_{st}^{(v)}\}$  is the vector of potentials and  $S$  is the net flow vector. Accordingly,  $S$  has 0 entries at each position except  $s$  and  $t$  where they are 1 and  $-1$  respectively. The problem of computing the flow on each edge consists in finding the potential vector  $V$  in (10). From a technical point of view, since only differences between potentials are considered, adding a constant to each potential will not affect the realizability of the system, which has actually an infinity of solutions. Indeed, 0 is always an eigenvalue of  $L$  and an associated eigenvector is the vector  $V_1$  whose all of its components are 1. Accordingly, the matrix  $L$  is not invertible. A way to handle this problem and to find a solution in (10) is to consider a similar system by adding 1 to each entry of  $L$ , as used by Klein and Randić [14]. Other methods can be used [12, 16] to solve (10). The obtained system will then have a unique solution that will also be a solution to the original problem, since the sum of the lines of matrices  $L$  and  $S$  are both equal to 0.

From a physical point of view, the equivalent resistance  $r_{st}$  between  $s$  and  $t$  is obtained by (7):

$$d_{st}^{(s)} - d_{st}^{(t)} = r_{st} \times i_{st} = r_{st} \quad (11)$$

since the global intensity  $i_{st}$  from  $s$  to  $t$  is  $1A$ .

Now, since we have all tools, we can define the Kirchhoff index [14]: it is the sum of equivalent resistances among all pairs of vertices in a graph  $G$ , i.e.

$$Kf(G) = \frac{1}{2} \sum_{s \in V} \sum_{t \in V} r_{st} = \sum_{s < t \in V} r_{st}. \quad (12)$$

Notice that the equivalent resistance may also be computed using the power dissipation law, which is described by the following general equation:

$$\text{Dissipated Power} = \text{Potential Difference} \times \text{Intensity}. \quad (13)$$

The dissipated power  $P_{st}^{uv}$  on an edge  $(u, v)$  between  $s$  and  $t$  is rewritten:

$$P_{st}^{uv} = (d_{st}^{(u)} - d_{st}^{(v)}) \times i_{st}^{uv} \stackrel{\text{by (8)}}{=} (d_{st}^{(u)} - d_{st}^{(v)})^2. \quad (14)$$

Since the global dissipated power  $P_{st}$  between  $s$  and  $t$  is the sum of dissipated powers on each edge, this global dissipated power is then

$$P_{st} = \sum_{(u,v) \in E} (d_{st}^{(u)} - d_{st}^{(v)})^2. \quad (15)$$

From a high point of view, the latter is also, by (13),

$$P_{st} = (d_{st}^{(s)} - d_{st}^{(t)}) \times i_{st} \stackrel{i_{st}=1}{=} d_{st}^{(s)} - d_{st}^{(t)} \stackrel{\text{by (11)}}{=} r_{st}, \quad (16)$$

thus the equivalent resistance  $r_{st}$  between  $s$  and  $t$  is defined by  $d_{st}^{(s)} - d_{st}^{(t)}$ , as well as

$$\sum_{(u,v) \in E} (d_{st}^{(u)} - d_{st}^{(v)})^2.$$

### 3 The expected distance and the RW index

Based upon the same model as the electric one, when viewing the network from the random walks point of view, we define *the expected length*  $\tilde{d}_{st}$  of the walk from  $s$  to  $t$  as the sum of the weighted lengths of paths from  $s$  to  $t$ . Observe that the weight of a path is its probability, i.e. the absolute value of its intensity. Moreover, the expected length  $\tilde{d}_{st}$  can be decomposed by the edges, i.e.

$$\tilde{d}_{st} = \sum_{(u,v) \in E} |d_{st}^{(u)} - d_{st}^{(v)}|. \quad (17)$$

As expected, the expected distance defined by the expected length between any pair of vertices is actually a distance, as established in the following theorem.

**Theorem 1** *Let  $G = (V, E)$  be a graph. The expected distance is a distance on  $V$ .*

*Proof* By definition in (17), the expected distance is well-defined. Reversing the source and the sink in a flow is imposed to take opposite potentials. Since the expected distance is defined by means of the absolute value of difference of potentials, it is symmetric. Separating distance is ensured because the only case where all potentials are equal is when the source are the sink.

The last property to check is the triangle inequality. Let  $s, r, t$  be three vertices in  $G$ . We need to prove that  $\tilde{d}_{st} \leq \tilde{d}_{sr} + \tilde{d}_{rt}$ . By the electric property, we know that if  $f_1$ , respectively  $f_2$ , is a flow through the network  $G$  from  $s$  to  $r$  with potentials  $d_{sr}^{(\cdot)}$  on  $V$ , respectively from  $r$  to  $t$  with potentials  $d_{rt}^{(\cdot)}$ , then  $f = f_1 + f_2$  is a flow from  $s$  to  $t$  with potentials

$$d_{st}^{(\cdot)} = d_{sr}^{(\cdot)} + d_{rt}^{(\cdot)}. \quad (18)$$

Accordingly, we have

$$\begin{aligned} \tilde{d}_{st} &= \sum_{(u,v) \in E} |d_{st}^{(u)} - d_{st}^{(v)}| \\ &\stackrel{(18)}{=} \sum_{(u,v) \in E} |(d_{sr}^{(u)} + d_{rt}^{(u)}) - (d_{sr}^{(v)} + d_{rt}^{(v)})| \\ &= \sum_{(u,v) \in E} |(d_{sr}^{(u)} - d_{sr}^{(v)}) + (d_{rt}^{(u)} - d_{rt}^{(v)})| \\ &\leq \sum_{(u,v) \in E} |d_{sr}^{(u)} - d_{sr}^{(v)}| + \sum_{(u,v) \in E} |d_{rt}^{(u)} - d_{rt}^{(v)}| \\ &= \tilde{d}_{sr} + \tilde{d}_{rt}. \end{aligned}$$

□

Similarly to the Wiener and the Kirchhoff indices, we define the RW index of a graph  $G$  as:

$$RW(G) = \frac{1}{2} \sum_{s \in V} \sum_{t \in V} \tilde{d}_{st} = \sum_{s < t \in V} \tilde{d}_{st}. \quad (19)$$

#### 4 Comparison of the Wiener, Kirchhoff, and RW indices

The following theorem establishes a comparison of the Wiener, Kirchhoff, and RW indices for all graphs. In addition, it yields that these three indices are the same only in the class of trees.

**Theorem 2** Let  $G = (V, E)$  be a graph, then we have:

$$Kf(G) \leq W(G) \leq RW(G). \quad (20)$$

Moreover, the following statements are equivalent:

- (i)  $Kf(G) = W(G)$ ,
- (ii)  $W(G) = RW(G)$ ,
- (iii)  $Kf(G) = RW(G)$ ,
- (iv)  $G$  is a tree.

*Proof* To be self-contained, we proof the first inequality. Let  $s$  and  $t$  be two vertices of  $G$ . From an electric point of view, the equivalent resistance  $r_{st}$  is always the length  $d_{st}$  of a shortest path from  $s$  to  $t$  when there is only one path between  $s$  and  $t$ , otherwise the equivalent resistance is strictly smaller than this value  $d_{st}$ . Then,  $r_{st} \leq d_{st}$  for every pair of vertices  $s, t \in V$ . Accordingly,  $Kf(G) \leq W(G)$ .

Now, we show the second inequality. Since the expected length  $\tilde{d}_{st}$  is by definition a convex combination of lengths of paths between  $s$  and  $t$ ,  $\tilde{d}_{st}$  is always greater than the smallest length of these paths, i.e.  $\tilde{d}_{st} \geq d_{st}$ . Therefore,  $W(G) \leq RW(G)$ .

We focus on the equivalences. First, we assume that  $G$  is a tree. By definition, there exists only one path between every pair of vertices. Thus,  $r_{st} = d_{st} = \tilde{d}_{st}$  for every pair of vertices  $s$  and  $t$ . Accordingly,

$$Kf(G) = W(G) = RW(G).$$

The last case is when  $G$  contains a cycle. Let  $C$  be an arbitrary cycle in  $G$  with the ordered vertices  $v_1, \dots, v_k$ . According to the Kirchhoff's circuit laws, assuming that  $v_{k+1} = v_1$ ,

$$\sum_{i=1}^k i_{v_1 v_2}^{v_i v_{i+1}} = 0,$$

i.e.

$$i_{v_1 v_2}^{v_1 v_2} = \sum_{i=2}^k i_{v_1 v_2}^{v_i v_{i+1}}.$$

Moreover,  $i_{v_1 v_2}^{v_1 v_2} \neq 0$  since the edge  $(v_1, v_2)$  is the shortest path between  $v_1$  and  $v_2$ . Accordingly,  $\sum_{i=2}^k i_{v_1 v_2}^{v_i v_{i+1}} \neq 0$ , i.e. there exists an edge  $(u, v)$  such that  $i_{v_1 v_2}^{uv} \neq 0$ . It is necessary that this edge is from a path from  $v_1$  to  $v_2$  with non-zero probability. Notice that the length of this path must be greater than  $d_{v_1 v_2} = 1$ . Thus, there are 2 paths from  $v_1$  to  $v_2$  with a non-zero probability and with different length, so  $d_{v_1 v_2} < \tilde{d}_{v_1 v_2}$ , which implies that  $W(G) < RW(G)$ , since  $d_{st} \leq \tilde{d}_{st} \forall s, t \in V$ . Moreover, by the same argument, we have that  $r_{v_1 v_2} < d_{v_1 v_2}$ . Because for every pair of vertices  $s, t \in V$ ,  $r_{st} \leq d_{st}$ , we conclude that  $Kf(G) < W(G)$  and then  $Kf(G) < RW(G)$ .  $\square$



## 5 Vertex and edge decompositions of Wiener, Kirchhoff, and RW indices

### 5.1 Decompositions of the Wiener index

The computation of the Wiener index may be achieved through two distinct, but related, concepts, that are the transmission and the edge betweenness [1, 9, 11]. Recall that the edge betweenness of an edge  $e$  is defined [11] as the number of shortest paths between pairs of vertices that use  $e$ . On the one hand, the transmission  $T_v$  of the vertex  $v$  is defined as:

$$T_v = \sum_{u \in V} d_{uv}. \quad (21)$$

The vertex decomposition of the Wiener index is naturally:

$$W(G) = \frac{1}{2} \sum_{v \in V} T_v. \quad (22)$$

On the other hand, if  $b_{uv}$  denotes the edge betweenness associated to the edge  $(u, v)$ , then Caporossi et al. [4] proved the edge decomposition of the Wiener index, i.e.

$$W(G) = \frac{1}{2} \sum_{u \in V} \sum_{v: (u,v) \in E} b_{uv} = \sum_{(u,v) \in E: u < v} b_{uv}. \quad (23)$$

### 5.2 Decompositions of the Kirchhoff index

Similarly to the decompositions of the Wiener index through transmission and edge betweenness, one can also decompose the Kirchhoff index through the related concepts of equivalent resistance and power dissipation. As stated by Klein and Randić [14], the equivalent resistance between any pair of vertices may be considered as a distance measure. It is then possible to compute the resistance transmission  $RT_v$  of a vertex  $v$  as follows:

$$RT_v = \sum_{u \in V} r_{uv}. \quad (24)$$

The Kirchhoff index then writes by the following vertex decomposition:

$$Kf(G) = \frac{1}{2} \sum_{v \in V} RT_v \quad (25)$$

If  $P^{uv}$  the power dissipation associated to each edge  $(u, v)$ , independently of the source and the sink, is defined by:

$$P^{uv} = \sum_{s, t \in V} \left( d_{st}^{(u)} - d_{st}^{(v)} \right)^2, \quad (26)$$

then the edge decomposition of the Kirchhoff index of a graph  $G$  is:

$$Kf(G) = \frac{1}{2} \sum_{u \in V} \sum_{v: (u,v) \in E} P^{uv} = \sum_{(u,v) \in E: u < v} P^{uv}. \quad (27)$$

### 5.3 Decompositions of the RW index

In the same way as the Wiener and the Kirchhoff indices, the RW index can be decomposed according to vertices and edges. As the transmission is defined by the use of shortest paths, the random walks transmission  $RWT_v$  of a vertex  $v$  is defined as:

$$RWT_v = \sum_{u \in V} \tilde{d}_{uv}. \quad (28)$$

The vertex decomposition of the RW index of a graph  $G$  is then

$$RW(G) = \sum_{v \in V} RWT_v. \quad (29)$$

Moreover, the edge decomposition of the Wiener and the Kirchhoff indices can be also extended to the RW index. Indeed, while the edge betweenness represents the number of shortest paths through pairs of vertices that uses a given edge, the random edge betweenness  $rb_{uv}$  of the edge  $(u, v)$  represents the sum of the probabilities that the edge  $(u, v)$  is used over all pairs of vertices as follows:

$$rb_{uv} = \sum_{s, t \in V} |d_{st}^{(u)} - d_{st}^{(v)}|. \quad (30)$$

Accordingly, the RW index of a graph  $G$  has the following edge decomposition:

$$RW(G) = \frac{1}{2} \sum_{u \in V} \sum_{v: (u,v) \in E} rb_{uv} = \sum_{(u,v) \in E: u < v} rb_{uv}. \quad (31)$$

### 5.4 Further remarks

As the adjusted<sup>1</sup> vertex betweenness [4], a vertex related value, was defined by

$$b'_v = \sum_{u: (u,v) \in E} b_{uv}, \quad (32)$$

---

<sup>1</sup> We are interested by the *adjusted* vertex betweenness, a centrality measure computing, for a given vertex  $u$ , the number of pairs of vertices such that  $u$  belongs to the shortest path between this pair of vertices. This definition allows that  $u$  can be an extremity of the shortest path, at the opposite of the vertex betweenness defined by Freeman [9,10].

we can define, in the same way, the vertex power dissipation  $p_v$  and the vertex random betweenness  $rb_v$  of a vertex  $v$  by the following equations:

$$p_v = \sum_{u:(u,v) \in E} p_{uv}, \quad (33)$$

$$rb_v = \sum_{u:(u,v) \in E} rb_{uv}. \quad (34)$$

## 6 Bounds and values of the RW index

### 6.1 Value of RW index for some families of graphs

In this section, we study the RW index for some families of graphs: paths, cycles, complete graphs and stars. Recall that a star is a tree on  $n$  vertices with a vertex of degree  $n - 1$ .

**Theorem 3** *We denote by*

- $C_n$  a cycle on  $n$  vertices,
- $P_n$  a path on  $n$  vertices,
- $S_n$  a star on  $n$  vertices,
- $K_n$  a complete graphs on  $n$  vertices.

*Then, the following results yield:*

- (i)  $RW(C_n) = RW(P_n) = \frac{n}{6}(n^2 - 1)$ ,
- (ii)  $RW(S_n) = RW(K_n) = (n - 1)^2$ .

*Proof* As mentionned in Theorem 2, in the class of trees, the three indices match. Moreover, Entringer, Jackson and Snyder [7] proved that  $W(P_n) = \frac{n}{6}(n^2 - 1)$  and  $W(S_n) = (n - 1)^2$ , proving two of the four equalities.

We consider the cycle  $C_n$  on  $n$  vertices. Since every vertex has the same role in the graph, we can fix one of them, say  $s$ , as a source. Let  $k$  be the distance between the source and the sink. By the property of the cycle, there are only two paths between  $s$  and  $t$ : one of length  $k$  and another one of length  $n - k$ . So, the potential difference all along each path is constant and inversely proportional to its length, i.e. the shortest path has a probability of  $\frac{n-k}{n}$  and

the longest one  $\frac{k}{n}$ . Accordingly,

$$\begin{aligned}
 RW(C_n) &= \frac{1}{2} \sum_{s \in V} \sum_{t \in V} \tilde{d}_{st} \\
 &= \frac{n}{2} \sum_{k=1}^{n-1} \left( k \frac{n-k}{n} + (n-k) \frac{k}{n} \right) \\
 &= \frac{n}{2} \sum_{k=1}^{n-1} \frac{2}{n} (k(n-k)) \\
 &= \sum_{k=1}^{n-1} (k(n-k)) \\
 &= \frac{n}{6} (n^2 - 1).
 \end{aligned}$$

The last case is the complete graph  $K_n$ . We fix arbitrarily a vertex  $s$  as a source and another vertex  $t$  as a sink, since every pair of vertices has the same feature. Paths of length strictly greater than 2 receive 0 as probability because these paths contain at least two successive vertices, distinct from the source and the sink, and because these two vertices have an identical role, and so the same potential. Thus, paths with non-zero probability have length 1 or 2. By a similar argument of the previous case, there are one path of length 1 with probability  $\frac{2}{n}$  and  $n-2$  paths of length 2 with probability  $\frac{1}{n}$ . Therefore,

$$\begin{aligned}
 RW(K_n) &= \frac{1}{2} \sum_{s \in V} \sum_{t \in V} \tilde{d}_{st} \\
 &= \frac{1}{2} n(n-1) \left( \frac{2}{n} + (n-2) \frac{2}{n} \right) \\
 &= (n-1)^2.
 \end{aligned}$$

□

## 6.2 Bounds on RW index

We observe that in the class of trees, stars and paths are extremal graphs of the RW index, since it is the case for the Wiener index [7]. More generally, we present the two following conjectures for extremal graphs, found by AutoGraphiX [2,3].

*Conjecture 1* The graph on  $n$  vertices that minimizes the RW index is either the star  $S_n$  or the complete graph  $K_n$ .

*Conjecture 2* The graph on  $n$  vertices that maximizes the RW index is either the path  $P_n$  or the cycle  $C_n$ .

If previous conjectures are true, we deduce the following conjecture by Theorem 3.

*Conjecture 3* Let  $G$  be a graph. Then,

$$(n-1)^2 \leq RW(G) \leq \frac{n}{6}(n^2-1).$$

Since the intensity of the edges are also related to the number of spanning trees [13,15], it is interesting to observe that the graphs on  $n$  vertices minimizing the number of spanning trees [8] are exactly the ones minimizing the RW index, according to the Conjecture 1, and that is not the case for the Kirchhoff index.

## 7 Concluding remarks and extensions

A possible extension of the RW index could be on graphs with weights on the edges. Indeed, in this paper, we consider the case when each edge is associated to a resistor of  $1\Omega$ . In the case of weighted graphs, we could set a resistor of  $w_e\Omega$  on each edge  $e$  with the weight  $w_e$ .

## References

1. J. M. Anthonisse, The rush in a directed graph, *Stichting Mathematisch Centrum. Mathematische Besliskunde*, 1–10, (1971).
2. G. Caporossi, Variable neighborhood search for extremal vertices: The system AutoGraphX-III, Technical Report G-2015-09, Groupe d'Études et de Recherche en Analyse des Décisions, (2015).
3. G. Caporossi and P. Hansen, Variable neighborhood search for extremal graphs: 1 The autographix system, *Discrete Mathematics*, 212(1):29–44, (2000).
4. G. Caporossi, M. Paiva, D. Vukićević, and M. Segatto, Centrality and betweenness: vertex and edge decomposition of the Wiener index, *MATCH-Communications in Mathematical and Computer Chemistry*, 68(1):293, (2012).
5. R. Diestel, Graph theory. 2005, *Grad. Texts in Math*, 101, (2005).
6. P.G. Doyle, J.L. Snell, Random walks and electric networks, *Free Software Foundation* (2000).
7. R. C. Entringer, D. E. Jackson, and D. A. Snyder, Distance in graphs, *Czechoslovak Mathematical Journal*, 26(2):283–296, (1976).
8. L. Feng, G. Yu, Z. Jiang, and L. Ren, Sharp upper bounds for the number of spanning trees of a graph, *Appl. Anal. Discrete Math.*, 2:255–259, (2008).
9. L. C. Freeman, A set of measures of centrality based on betweenness, *Sociometry*, 35–41, (1977).
10. L. C. Freeman, Centrality in social networks conceptual clarification, *Social Networks*, 1(3):215–239, (1979).
11. M. Girvan, M.E.J. Newman, Community structure in social and biological networks, *Proceedings of the national academy of sciences*, 99(12):7821–7826, (2002).
12. I. Gutman and W. Xiao, Generalized inverse of the laplacian matrix and some applications, *Bulletin de l'Académie Serbe des Sciences at des Arts (Cl. Math. Natur.)*, 129:15–23, (2004).
13. G. Kirchhoff, Ueber die auflösung der gleichungen, auf welche man bei der untersuchung der linearen vertheilung galvanischer ströme geführt wird, *Annalen der Physik*, 148(12):497–508, (1847).

14. D. J. Klein and M. Randić, Resistance distance, *Journal of Mathematical Chemistry*, 12(1):81–95, (1993).
15. R. Lyons and Y. Peres, *Probability on Trees and Networks*, Cambridge University Press, (2016). Available at <http://pages.iu.edu/~rdlyons/>.
16. M. E. J. Newman, A measure of betweenness centrality based on random walks, *Social Networks*, 27(1):39–54, (2005).
17. K. Stephenson and M. Zelen, Rethinking centrality: Methods and examples, *Social Networks*, 11(1):1–37, (1989).
18. H. Wiener, Structural determination of paraffin boiling points, *Journal of the American Chemical Society*, 69(1):17–20, (1947).