SIAM J. MATRIX ANAL. APPL.
Vol. 44, No. 2, pp. 648–669

AN EDGE CENTRALITY MEASURE BASED ON THE KEMENY CONSTANT*

DIEGO ALTAFINI[†], DARIO A. BINI[‡], VALERIO CUTINI[†], BEATRICE MEINI[‡], AND FEDERICO POLONI[§]

Abstract. A new measure $c(e)$ of the centrality of an edge e in an undirected graph G is introduced. It is based on the variation of the Kemeny constant of the graph after removing the edge e. The new measure is designed to satisfy certain monotonicity and positivity properties, and hence using it one can avoid the *Braess paradox*, i.e., the phenomenon in which removing an edge can increase the connectivity of a network rather than reduce it. A numerical method for computing $c(e)$ is introduced, and a regularization technique is designed in order to deal with cut-edges and disconnected graphs. Numerical experiments performed both on artificial examples and on real road networks show that this measure is particularly effective in revealing bottleneck roads whose removal would greatly reduce the connectivity of the network.

Key words. network theory, centrality measure, Kemeny constant, Markov chain

MSC codes. 65F30, 05C82, 91D30

DOI. 10.1137/22M1486728

1. Introduction.

In network analysis, several measures of the importance of nodes of a graph have been introduced, each with different modelistic meanings and mathematical formulations. For instance, in [2, 16] the communicability between two nodes i, j of a graph G is defined as the (i, j) entry in the exponential of the adjacency matrix of G. The exponential of a matrix is also at the basis of the definition of importance given in [15]. Other measures based on the computation of matrix functions are introduced in [4], where a parameterized node centrality measure is proposed, and in [3], where directed networks are analyzed.

Measures for the centrality of *edges*, rather than nodes, are less common but still frequent in the literature. Edge betweenness, i.e., counting how many shortest paths go through a certain edge, is suggested in [17]. Replacing shortest paths with random walks gives a measure that can also be interpreted in terms of resistance of electrical networks, called *Edge current flow betweenness centrality* [8]. More examples appear in [11, 12, 26, 27, 30]. Notably, in [14], the authors suggest the idea of considering the variation of the Kemeny constant when an edge is removed from a graph.

In this paper, following [14], we introduce a new definition of edge centrality based on a modified variation of the Kemeny constant and perform a theoretical and computational analysis.

Given a connected, weighted, undirected, finite graph $G = (V, E)$ with $n = |V|$ nodes, the random walk associated to it is a Markov chain with adjacency matrix $P = D^{-1}A$, where A is the adjacency matrix of G, $D = \mathrm{diag}(d)$, $d = A\mathbf{1}$. Here

*Received by the editors March 28, 2022; accepted for publication (in revised form) November 28, 2022; published electronically May 19, 2023.
https://doi.org/10.1137/22M1486728
Funding: This work has been partially supported by University of Pisa's project PRA_2020_61, and by GNCS of INdAM.
[†]Dipartimento di Ingegneria dell'Energia, dei Sistemi, del Territorio e delle Costruzioni, Università di Pisa, Italy (diego.altafini@phd.unipi.it, valerio.cutini@unipi.it).
[‡]Dipartimento di Matematica, Università di Pisa, Italy (bini@dm.unipi.it, beatrice.meini@unipi.it).
[§]Dipartimento di Informatica, Università di Pisa, Italy (federico.poloni@unipi.it).

$\mathbf{1} = (1, 1, \ldots, 1)^T$ is the vector of all ones, and $\mathrm{diag}(v)$ denotes the diagonal matrix having the entries of the vector v on the diagonal.

Let $\pi = (\pi_k)_{k=1,\ldots,n}$ be the invariant measure of said Markov chain, so that $\pi^T P = \pi^T$ and $\pi^T \mathbf{1} = 1$. The *Kemeny constant* $K(G)$ is defined as the expected first-passage time of the Markov chain from a predetermined state $i \in V$ to a state $j \in V$ drawn randomly according to the probability distribution π. It is a surprising but well-studied fact that this definition does not depend on i [23].

The Kemeny constant gives a global measure of the nonconnectivity of a network [5, 14, 28]. Indeed, if G is not connected, then the Kemeny constant cannot be defined, or, in different words, it takes the value infinity.

Following the idea of [14], we may define the Kemeny-based *centrality measure* of the edge $e \in E$ as

$$\widetilde{c}(e) := K((V, E \setminus \{e\})) - K((V, E)),$$

i.e., the change of the connectivity of the graph measured by the Kemeny constant, when the edge e is removed from the graph.

In matrix form, the value of $\widetilde{c}(e)$ can be given in terms of the eigenvalues of the Laplacian matrix $D - A$ and of a suitable rank 2 correction. For the symmetry of the Laplacian, the value of $\widetilde{c}(e)$ can be also expressed in terms of the eigenvalues of the symmetric matrix $Y = D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$ and of the eigenvalues of $Y + C$, where C is a symmetric correction of rank 2.

A surprising feature of this definition of centrality measure is that there exist graphs where $\widetilde{c}(e)$ is negative for some e; an elementary example is shown in section 6.1. Hence one arrives at the paradoxical conclusion that removing an edge *increases* the connectivity of a network rather than reduces it. In the literature, this fact is known as the Braess paradox [7], [14], and it has been observed for several connectivity measures and metrics on road networks. Nevertheless, this phenomenon makes it more difficult to give interpretations of the measure in terms of ranking and comparing edges.

Hence, we propose a modified centrality measure $c(e)$, which is always nonnegative for any graph and for any edge e. The underlying idea consists in replacing the correction C with a positive semidefinite matrix of rank 1. From the model point of view, the new correction consists in replacing the edge $e = \{i, j\}$ with two loops $\{i, i\}$ and $\{j, j\}$ with the same weight. The new definition not only ensures the desirable nonnegativity property $c(e) \geq 0$ but is also faster to compute, since it is based on a rank 1 modification of the adjacency matrix rather than a rank 2 one.

These definitions cannot be applied in the case where e is a cut-edge, i.e., $(V, E \setminus \{e\})$ is not connected; in fact, in this case, the definition would yield $\widetilde{c}(e) = c(e) = \infty$. To overcome this drawback, we introduce the concept of a *regularized centrality measure* $c_r(e)$, depending on a regularization parameter $r > 0$. The idea is to replace the Laplacian matrix $D - A$ with the regularized Laplacian matrix $(1 + r)D - A$ in the formulas that give the Kemeny constant. The regularization parameter can be interpreted in terms of the teleportation probability in the PageRank model. If e is not a cut-edge, then $\lim_{r \to 0} c_r(e) = c(e)$; otherwise, if e is a cut-edge, then $\lim_{r \to 0} c_r(e) = \infty$. Moreover, we show that, if e is a cut-edge, then the quantity $r^{-1} - c_r(e)$ is nonnegative and has a finite limit for $r \to 0$; this suggests the following definition of a *filtered* Kemeny-based centrality measure:

$$c_r^F(e) := \begin{cases} c_r(e) & \text{if } e \text{ is not a cut-edge,} \\ r^{-1} - c_r(e) & \text{if } e \text{ is a cut-edge.} \end{cases}$$

The modified measure defined in this way is still nonnegative and seems particularly effective in highlighting bottlenecks in road networks, or so-called *weak ties* [20] that bridge different clusters. Identifying weak ties has potential applications in other fields such as the analysis of social, economical, and biological networks where this measure has the potential to reveal key connections that may go unnoticed under other metrics [20]. In particular, in road-circulation networks, configurational metrics based on closeness and betweenness centralities have a long-standing tradition in addressing the potential movement patterns. Nevertheless, these metrics are limited in their explanation of the relative importance of the road-network elements within a system, an aspect that has received meager attention [1]. The proposed measure aims to highlight the relative importance of road-elements that have a crucial bridge function (the *weak ties*) and their role in connecting clusters of road elements (the *strong ties*). Moreover, it demonstrates the overall characteristics of redundancy within a system, identifying the roads that, if interrupted, may cause certain parts of the network to collapse. These uses open concrete perspectives into urban and regional planning applications related to fragility analysis such as: emergency routing, disaster risk prevention, and risk mitigation.

We provide efficient algorithms implementing the computation of the measure either of a single edge or of all the edges of a graph. The main tools in the algorithm design are the Sherman–Woodbury–Morrison formula and the Cholesky factorization of the regularized Laplacian matrix $(1+r)D-A$.

Our algorithms have been tested both on artificial examples and on graphs representing real road networks; in particular, we have considered the maps of Pisa and of the entire Italian region Tuscany. Our study has been motivated by an application in which it is important to identify weak ties in a road network; hence the measure is particularly useful in this context.

From our numerical experiments, reported in the paper, it turned out that this measure is robust, effective, and realistic from the model point of view. Moreover, its computation is sufficiently fast even for large road networks. Comparisons with other centrality measures from [16] have been performed. It turns out that our model, unlike the ones based on PageRank and Betweenness of the dual graph, succeeds in detecting bridges on the river Arno and overpasses over the railroad line as important bottleneck roads, or weak ties, in the Pisa road map. Edge betweenness and Edge current-flow betweenness are the only two measures (among those considered) that succeed in the same task, even though they succeed only partially and at a much higher CPU time. The time required by the other betweenness-based measures on planar networks of roads is comparable with that required by the Kemeny-based measure. More details concerning applications of the Kemeny-based centrality measure to road networks can be found in [1].

The paper is organized as follows. In section 2 we recall some properties of the Kemeny constant. In section 3 the Kemeny-based centrality measure is introduced and a matrix analysis is performed, while in section 4 a modified definition is proposed in order to always return non-negative results. The regularized and filtered centrality measures are proposed in section 5. Section 6 contains further theoretical developments concerning the above centrality measures. Section 7 is devoted to computational aspects and numerical experiments. Conclusions are drawn in section 8.

2. The Kemeny constant. Let P be the $n\times n$ transition matrix of an irreducible finite Markov chain, and let π be its steady state vector. Denote by $K(P)$ the Kemeny constant of P; more formally,

$$K(P) = \sum_j m_{ij}\pi_j,$$

where m_{ij} is the expected first-passage time from node i to node j, and π is the invariant measure of P. We recall some properties which allow us to express the Kemeny constant in terms of the trace of a suitable matrix. Such expressions will be useful in the analysis performed in the next sections.

LEMMA 2.1 ([25, section 3]). *Let $g, h \in \mathbb{R}^n$ be column vectors with $h^T g = 1$, $h^T\mathbf{1} \neq 0$, $\pi^T g \neq 0$. Then, the inverse $Z := (I - P + gh^T)^{-1}$ exists, and*

$$K(P) = \operatorname{trace}(Z) - \pi^T Z\mathbf{1},$$

independently of g, h.

By setting $g = \mathbf{1}$, one gets the following corollary.

COROLLARY 2.2. *Let h be a column vector with $h^T\mathbf{1} = 1$; then, $Z = (I - P + \mathbf{1}h^T)^{-1}$ exists, and*

$$K(P) = \operatorname{trace}(Z) - 1. \tag{2.1}$$

Proof. Setting $g = \mathbf{1}$, we have $Z^{-1}\mathbf{1} = (I - P)\mathbf{1} + \mathbf{1}(h^T\mathbf{1}) = 0 + \mathbf{1} = \mathbf{1}$; hence we can simplify the second part of the formula. □

Since P is an irreducible stochastic matrix, it has a simple eigenvalue equal to 1. The Kemeny constant can be expressed by means of the eigenvalues different from 1, according to the following result.

COROLLARY 2.3. *Let $\lambda_1 = 1, \lambda_2, \ldots, \lambda_n$ be the spectrum of P. Then,*

$$K(P) = \sum_{\ell=2}^{n} \frac{1}{1 - \lambda_\ell}. \tag{2.2}$$

Proof. Take a Jordan form $P = WJW^{-1}$ with $W_{:,1} = \mathbf{1}$, $W^{-1}_{1,:} = \pi^T$, and $\operatorname{diag}(J) = (1, \lambda_2, \lambda_3, \ldots, \lambda_n)$ (reordering $\lambda_2, \ldots, \lambda_n$ if necessary). Then, one has

$$I - P + \mathbf{1}\pi^T = W(I - J + e_1 e_1^T)W^{-1} = WTW^{-1},$$

where T is upper triangular with $\operatorname{diag}(T) = (1, 1 - \lambda_2, 1 - \lambda_3, \ldots, 1 - \lambda_n)$. Plugging this expression into (2.1), we get

$$K(P) = \operatorname{trace}(Z) - 1 = \operatorname{trace}(WT^{-1}W^{-1}) - 1 = \operatorname{trace}(T^{-1}) - 1 = \sum_{\ell=2}^{n} \frac{1}{1 - \lambda_\ell}.$$ □

3. A centrality measure based on the Kemeny constant. We open this section by defining some standard terminology in graph theory, following for instance [22, Chapter 39]. An (undirected) graph (sometimes also called *network*) is a pair $G = (V, E)$, where V is a finite set of *nodes* of cardinality $|V| = n$, and E is a set of *edges*, i.e., pairs $\{i, j\} \subset V$, of cardinality $|E| = m$. A *path* (from $i_0 \in V$ to $i_\ell \in V$ of length ℓ) is a sequence of distinct nodes $i_0, i_1, i_2, \ldots, i_\ell$ connected by edges, i.e., $\{i_{k-1}, i_k\} \in E$ for each $k = 1, 2, \ldots, \ell$. A graph is called *connected* if for every pair of nodes $i, j \in V$, $i \neq j$, there is a path from i to j. A graph is called *weighted* if it is equipped with a function $a : E \to \mathbb{R}_+$ that associates a nonnegative *weight* to each edge. An edge $e \in E$ is called a *cut-edge* if $G = (V, E)$ is connected, but the graph $\widehat{G} = (V, E \setminus \{e\})$ obtained by removing said edge is disconnected.

We assume that we start from a weighted, connected, undirected graph $G = (V, E)$ containing no *loops*, i.e., edges $\{i, j\}$ with $i = j$. Some of the procedures that we describe in the following, however, will introduce modifications to this graph that create loops or lose connectedness.

The *adjacency matrix* of a graph is the $n \times n$ symmetric matrix $A = (a_{ij})$ such that a_{ij} is equal to the weight of edge $\{i, j\}$, if said edge exists in E, or zero otherwise. To a graph we can associate the Markov chain with transition matrix $P = D^{-1}A$, where $D = \operatorname{diag}(d)$, $d = A\mathbf{1}$. We can then define its Kemeny constant $K(G) := K(P)$. The Kemeny constant gives a global measure of the connectivity of a network: small values of the constant correspond to highly connected networks, and large values correspond to a low connectivity.

To obtain a relative measure that takes into account the importance of each edge $e = \{i, j\} \in E$, we can define the *Kemeny-based centrality measure* as

$$\widetilde{c}(e) := K((V, E \setminus \{e\})) - K((V, E)), \tag{3.1}$$

i.e., the change in K obtained by removing the edge e. This quantity is well defined assuming that e is not a cut-edge.

Removing one edge $e = \{i, j\}$ corresponds to zeroing out the entries $a_{i,j}$ and $a_{j,i}$. This leads to the new adjacency matrix

$$\widehat{A} = A - a_{i,j} U \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} U^T, \quad U = \begin{bmatrix} e_i & e_j \end{bmatrix} \in \mathbb{R}^{n \times 2}, \tag{3.2}$$

where e_i and e_j are the ith and the jth columns of the identity matrix I, respectively. This removal changes the transition matrix P into the matrix $\widehat{P} = \widehat{D}^{-1}\widehat{A}$, where $\widehat{D} = \operatorname{diag}(\hat{d})$, $\hat{d} = \widehat{A}\mathbf{1}$, which differs from P only in rows i and j since $\hat{d} = d - a_{i,j}(e_i + e_j)$. Hence we have

$$\widehat{P} = P + UV^T, \tag{3.3}$$

where

$$V^T = \begin{bmatrix} s_i & 0 \\ 0 & s_j \end{bmatrix} U^T A - a_{i,j} \begin{bmatrix} 0 & (d_i - a_{i,j})^{-1} \\ (d_j - a_{i,j})^{-1} & 0 \end{bmatrix} U^T, \tag{3.4}$$

with $s_i = \frac{a_{i,j}}{d_i(d_i - a_{i,j})}$, $s_j = \frac{a_{i,j}}{d_j(d_j - a_{i,j})}$.

THEOREM 3.1. *Let G be a connected weighted graph, and suppose edge e is not a cut-edge. Then, for the centrality measure defined in* (3.1) *we have*

$$\widetilde{c}(e) = \operatorname{trace}((I - V^T Z U)^{-1} V^T Z^2 U), \tag{3.5}$$

where $U = \begin{bmatrix} e_i & e_j \end{bmatrix}$, V^T is defined in (3.4), *and $Z = (I - P + \mathbf{1}h^T)^{-1}$ is as in Corollary* 2.2.

Proof. We have

$$\begin{aligned} \hat{Z} &:= (I - \hat{P} + \mathbf{1}h^T)^{-1} \\ &= (I - P + \mathbf{1}h^T - UV^T)^{-1} \\ &= Z + ZU(I - V^T Z U)^{-1} V^T Z, \end{aligned}$$

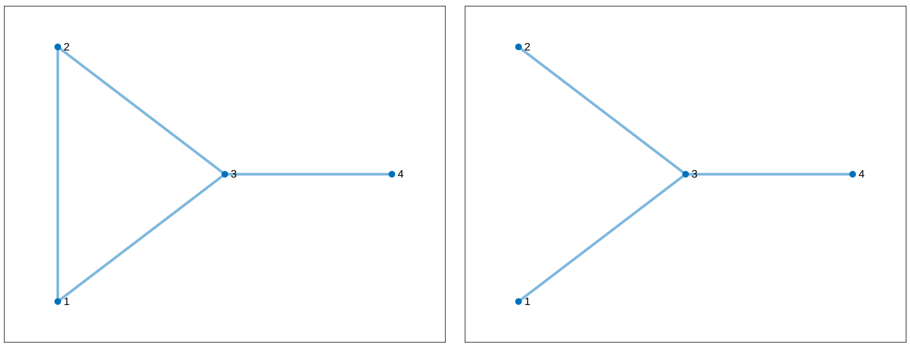

FIG. 4.1. *The graph on the right is obtained from that on the left by removing the edge* $\{1,2\}$. *The two graphs have Kemeny constants* $\frac{61}{24} \approx 2.54$ *and* 2.5, *respectively.*

where we have used (3.3) and in the last step the Sherman–Woodbury–Morrison matrix identity [19, section 2.1.4]. We now use (2.1) and write

$$\begin{aligned}\widetilde{c}(e) &= K(\hat{P}) - K(P) = \mathrm{trace}(\hat{Z}) - \mathrm{trace}(Z) = \mathrm{trace}(\hat{Z} - Z)\\ &= \mathrm{trace}(ZU(I - V^T ZU)^{-1} V^T Z)\\ &= \mathrm{trace}((I - V^T ZU)^{-1} V^T Z^2 U),\end{aligned}$$

using the identity $\mathrm{trace}(MN) = \mathrm{trace}(NM)$ [22, section 69.2, Fact 2]. ◻

Theorem 3.1 allows us to compute the centrality measure of one edge at essentially the cost of applying the matrix Z to four vectors.

4. A nonnegative Kemeny-based centrality measure. Intuitively, one expects that the connectivity of a graph should not increase if an edge is removed from the graph. Therefore, if the Kemeny constant properly describes the nonconnectivity of a graph, then it should not decrease if an edge is removed. In terms of definition of centrality measure given in (3.1), we expect that $\widetilde{c}(e) \geq 0$. Unfortunately, it is not so.

In fact, there are cases where the Kemeny constant of a graph can decrease if an edge is removed, like in the graph with edges $E = \{\{1,2\},\{1,3\},\{2,3\},\{3,4\}\}$, shown in Figure 4.1, on the left. Its Kemeny constant is $\frac{61}{24} \approx 2.54$. Removing the edge $\{1,2\}$, we get the graph on the right, which has a smaller Kemeny constant, i.e., 2.5. That is, the centrality measure of the edge $\{1,2\}$ in this graph is *negative*. This fact is known in the literature as the Braess paradox [14, 7].

In order to overcome this odd behavior of the model, where the measure $\widetilde{c}(e)$ can take negative values, we propose a simple modification which also makes the computation of the centrality measure a less expensive task.

Observe that removing the edge $\{i,j\}$ from the graph consists in performing a correction to the adjacency matrix A of rank 2 in order to obtain the new matrix $\widehat{A}$; compare with (3.2). This correction is such that the vector $d = A\mathbf{1}$ differs from the vector $\hat{d} = \widehat{A}\mathbf{1}$ in the components i and j. On the other hand, defining $\widehat{A}$ in a different way, by means of the expression

$$\widehat{A} = A + a_{i,j} vv^T, \quad v = e_i - e_j, \tag{4.1}$$

has the effect of zeroing the entries $a_{i,j}$ and $a_{j,i}$ in A, and of adding $a_{i,j}$ to the diagonal entries in positions (i,i) and (j,j). In terms of graph, this correction consists in removing the edge $\{i,j\}$ and adding the two loops $\{i,i\}$ and $\{j,j\}$ with the same weight $a_{i,j}$. More formally, we define

$$c(e) = K(\widehat{G}_e) - K(G), \tag{4.2}$$

where

$$\widehat{G}_e = (V, E \setminus \{e\} \cup \{\{i,i\},\{j,j\}\}), \quad \text{with } e = \{i,j\},$$

and the new weights in the adjacency matrix $\widehat{A}$ are $\widehat{a}_{ii} = \widehat{a}_{jj} = a_{ij}$.

4.1. A symmetrized formulation. In this section, we slightly modify the formulas for the Kemeny constant so that they involve symmetric matrices, borrowing from the idea of normalized Laplacian $D^{-1/2}(D-A)D^{-1/2} = I - D^{-1/2}AD^{-1/2}$ which is ubiquitous in graph theory. Note that the symmetrized Laplacian has been used to compute Kemeny constants, for instance, in [10].

Observe that $D^{\frac{1}{2}}PD^{-\frac{1}{2}} = D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ is a symmetric matrix having the same spectrum of P. Therefore, we may rewrite the expression in Corollary 2.2 to

$$K(P) = \operatorname{trace}(W) - 1, \quad W = D^{1/2}ZD^{-1/2} = (I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} + D^{\frac{1}{2}}\mathbf{1}h^T D^{-\frac{1}{2}})^{-1}.$$

We choose $h = \frac{1}{\sum_{i=1}^n d_i} d$, so that the last summand $D^{\frac{1}{2}}\mathbf{1}h^T D^{-\frac{1}{2}}$ is symmetric with an eigenvalue 1, and obtain

$$K(P) = \operatorname{trace}(W) - 1, \quad W = \left(I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} + \frac{1}{\|d\|_1} D^{\frac{1}{2}}\mathbf{1}\mathbf{1}^T D^{\frac{1}{2}}\right)^{-1}. \tag{4.3}$$

In the above expression, the matrix W is real symmetric.

In the same manner, for the modified network after the removal of an edge e we have

$$K(\widehat{P}) = \operatorname{trace}(\widehat{W}) - 1, \quad \widehat{W} = \left(I - \widehat{D}^{-\frac{1}{2}}\widehat{A}\widehat{D}^{-\frac{1}{2}} + \frac{1}{\|\widehat{d}\|_1} \widehat{D}^{\frac{1}{2}}\mathbf{1}\mathbf{1}^T \widehat{D}^{\frac{1}{2}}\right)^{-1}. \tag{4.4}$$

Thus, we may write $c(e) = \operatorname{trace}(\widehat{W} - W)$, where $\widehat{W} - W$ is a low rank symmetric matrix. This fact enables us to exploit results on the eigenvalues of symmetric matrices like the Courant–Fischer theorem [6, Chapter III] to prove certain monotonicity properties of the new measure.

LEMMA 4.1. *Let A_1, A_2, A_3 be real symmetric $n \times n$ matrices such that $A_3 = A_1 + A_2$, and let $\mu_i^{(1)}$, $\mu_i^{(2)}$, $\mu_i^{(3)}$, $i = 1, \ldots, n$, be their eigenvalues, respectively, ordered in nondecreasing order. Then $\mu_i^{(1)} + \mu_1^{(2)} \le \mu_i^{(3)} \le \mu_i^{(1)} + \mu_n^{(2)}$ for $i = 1, \ldots, n$.*

We are ready to prove the following result.

THEOREM 4.2. *Let A be the $n \times n$ adjacency matrix of an undirected weighted connected graph, let $i, j \in \{1, \ldots, n\}$ be such that the edge $e = \{i,j\}$ is not a cut-edge, and let $\widehat{A}$ be the adjacency matrix defined in* (4.1). *Then for the centrality measure defined as $c(e) = K(\widehat{P}) - K(P)$, we have $c(e) \ge 0$, where $P = D^{-1}A$, $\widehat{P} = \widehat{D}^{-1}\widehat{A}$, $D = \widehat{D} = \operatorname{diag}(d)$, $d = A\mathbf{1} = \widehat{A}\mathbf{1}$.*

Proof. Write $c(e)$ in terms of the symmetrized formulation according to (4.3) and (4.4), and get

$$c(e) = \sum_{\ell=2}^n \frac{1}{1 - \hat{\lambda}_\ell} - \sum_{\ell=2}^n \frac{1}{1 - \lambda_\ell}, \tag{4.5}$$

where $\hat\lambda_\ell$ and λ_ℓ, $\ell=1,\dots,n$, are the eigenvalues, sorted in nonincreasing order, of the symmetric matrices

$$\widehat H := \widehat D^{-\frac12}\widehat A\widehat D^{-\frac12} - \frac{1}{\|\hat d\|_1}\widehat D^{\frac12}\mathbf{1}\mathbf{1}^T\widehat D^{\frac12} \quad \text{and} \quad H := D^{-\frac12}AD^{-\frac12} - \frac{1}{\|d\|_1}D^{\frac12}\mathbf{1}\mathbf{1}^T D^{\frac12},$$

respectively. On the other hand, since $d=\hat d$ and $D=\widehat D$, we have $\widehat H = H + a_{i,j}D^{-\frac12}vv^TD^{-\frac12}$. The matrix $a_{i,j}D^{-\frac12}vv^TD^{-\frac12}$ has $n-1$ eigenvalues equal to 0 and one eigenvalue equal to $a_{i,j}v^TD^{-1}v = a_{i,j}(d_i^{-1}+d_j^{-1})$. Applying Lemma 4.1 with $A_1 = H$ and $A_2 = a_{i,j}D^{-\frac12}vv^TD^{-\frac12}$ yields the inequality $\lambda_\ell \le \hat\lambda_\ell \le \lambda_\ell + a_{i,j}(d_i^{-1}+d_j^{-1})$. This implies that $c(e)\ge 0$ in view of (4.5). ∎

Computing the value of $c(e)$ defined in (4.2) is cheaper than computing the quantity defined in (3.1). To this regard, we have the following result.

THEOREM 4.3. *Under the assumptions of Theorem* 4.2 *we have*

$$c(e) = a_{i,j}v^TD^{-\frac12}W\widehat W D^{-\frac12}v,$$

where $W^{-1} = I - D^{-\frac12}AD^{-\frac12} + \frac{1}{\|d\|_1}D^{\frac12}\mathbf{1}\mathbf{1}^TD^{\frac12}$, $\widehat W^{-1} = W^{-1} - a_{i,j}D^{-\frac12}vv^TD^{-\frac12}$. *Moreover,* $\widehat W = W - \tau a_{i,j}WD^{-\frac12}vv^TD^{-\frac12}W$ *for* $\tau = -1/(1-a_{i,j}v^TD^{-\frac12}WD^{-\frac12}v)$.

Proof. By using symmetrization, we have $c(e) = \mathrm{trace}(\widehat W - W) = \mathrm{trace}(\widehat W(W^{-1} - \widehat W^{-1})W)$. On the other hand, $W^{-1} - \widehat W^{-1} = a_{i,j}D^{-\frac12}vv^TD^{-\frac12}$, so that $c(e) = \mathrm{trace}(a_{i,j}\widehat W D^{-\frac12}vv^TD^{-\frac12}W) = a_{i,j}v^TD^{-\frac12}W\widehat WD^{-\frac12}v$. The expression for $\widehat W$ follows from the Sherman–Woodbury–Morrison identity. ∎

The above result can be used to obtain an effective expression for computing $c(e)$. To this end, rewrite W as

$$W = D^{\frac12}S^{-1}D^{\frac12}, \quad S = D - A + \frac{1}{\|d\|_1}dd^T$$

so that

$$\widehat W = D^{\frac12}\widehat S^{-1}D^{\frac12}, \quad \widehat S = D - \widehat A + \frac{1}{\|d\|_1}dd^T = S - a_{i,j}vv^T.$$

Moreover, from the Sherman–Woodbury–Morrison formula we have

$$\widehat S^{-1} = S^{-1} + \frac{a_{i,j}}{1-a_{i,j}v^TS^{-1}v}S^{-1}vv^TS^{-1}.$$

Whence in view of Theorem 4.3 we obtain

$$c(e) = a_{i,j}v^TS^{-1}D\widehat S^{-1}v = a_{i,j}v^TS^{-1}DS^{-1}v + \frac{a_{i,j}^2v^TS^{-1}v}{1-a_{i,j}v^TS^{-1}v}v^TS^{-1}DS^{-1}v.$$

From the above result we obtain the following representation of $c(e)$:

$$\begin{aligned} &c(e) = \frac{\beta}{1-\alpha}, \quad \alpha = a_{i,j}v^Tx = a_{i,j}(x_i - x_j), \quad \beta = a_{i,j}x^TDx,\\ &x = S^{-1}v, \quad S = D - A + \frac{1}{\|d\|_1}dd^T. \end{aligned} \tag{4.6}$$

Observe that α and β in (4.6) can be rewritten as

$$\begin{aligned} &\alpha = a_{i,j}(e_i-e_j)^TF(e_i-e_j), \quad \beta = a_{i,j}(e_i-e_j)^TQ(e_i-e_j),\\ &F = S^{-1}, \quad Q = FDF. \end{aligned} \tag{4.7}$$

Another observation is that the matrix S is positive definite since it is invertible and is the sum of two semidefinite matrices. Therefore, it admits the Cholesky factorization $S = LL^T$.

The major computational effort in computing $c(e)$ by means of (4.6) consists in solving the system $Sx = v$. If one has to compute the centrality measure of a single edge $\{i, j\}$, then two strategies can be designed for this task. A first possibility consists in computing the Cholesky factorization of S and solving the two triangular systems. This approach costs $O(n^3)$ arithmetic operations, as the dominating cost is the one of the Cholesky factorization. A second possibility consists in applying an iterative method for solving the linear system with matrix S that exploits the low cost of the matrix-vector product, say, Richardson iteration or the preconditioned conjugate gradient method. This approach costs $O(m+n)$ operations per iteration, where m is the number of nonzero entries of the adjacency matrix. Thus, it is cheaper than the former approach as long as the number of required iterations is less than n^3/m.

A different conclusion holds in the case where the centrality measures $c_{i,j}$ of all edges $e = \{i, j\}$ must be computed. In fact, in this case, the cost is $O(n^3 + m)$, by relying on the following computation that is based on (4.7):

1. Compute $F = S^{-1}$ and $Q = FDF$;
2. for all $i < j$ such that $a_{i,j} \neq 0$ compute:
   (a) $\alpha = a_{i,j}(f_{i,i} + f_{j,j} - 2f_{i,j})$,
   (b) $\beta = a_{i,j}(q_{i,i} + q_{j,j} - 2q_{i,j})$,
   (c) $c_{i,j} = \beta/(1-\alpha)$.

The overall cost of the above approach is dominated by the cost of step 1, i.e., $O(n^3)$ arithmetic operations. The drawback of this approach is that all the n^3 entries of the matrices F and Q must be stored. This can be an issue if n takes very large values.

Another issue is the potentially large condition number of the matrix S. The matrix $D^{-1}S = I - P + \mathbf{1}h^T$ has eigenvalues $1, 1-\lambda_2, \ldots, 1-\lambda_n$: they are equal to those of $I - P$, except that the correction $\mathbf{1}h^T$ has the effect of replacing the zero eigenvalue by $h^T\mathbf{1} = 1$. Hence we can expect ill-conditioning in S when $\lambda_2 \approx 1$, i.e., when the network is almost disconnected. When the network is disconnected, $\lambda_2 = 1$ and S becomes exactly singular. A way to overcome this difficulty consists in applying a sort of regularization in the inversion of the matrix S. This regularization, which is the subject of the next section, turns out to be helpful not only in cases when S is ill-conditioned but also in dealing with the cut-edges of the network.

5. Regularized Kemeny-based centrality measure. Let $r > 0$ be a (typically small) regularization parameter and, with the notation of the previous sections, define the regularized Kemeny constant as

$$\begin{aligned} K_r(G) &= \operatorname{trace}(((1+r)I - D^{-1}A + \mathbf{1}h^T)^{-1}) - (1+r)^{-1} \\ &= \operatorname{trace}\left(\left((1+r)I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} + \frac{1}{\|d\|_1}D^{\frac{1}{2}}\mathbf{1}\mathbf{1}^T D^{\frac{1}{2}}\right)^{-1}\right) - (1+r)^{-1}, \end{aligned}$$

where, for the second expression, we used the symmetrized version. Observe that, with respect to the standard definition, we have increased the diagonal entries of the matrix $W^{-1} = I - D^{-1}A + \mathbf{1}h^T$ by the quantity r. On one hand, this modification reduces the condition number of W; on the other hand, it allows us to deal with the situations where W is singular—for instance, in the case where the graph is not connected.

Since the eigenvalues of W^{-1} are $1, 1-\lambda_2, 1-\lambda_3, \ldots, 1-\lambda_n$, we have

$$K_r(G) = \operatorname{trace}(rI + W^{-1})^{-1} - \frac{1}{1+r} = \sum_{\ell=1}^{n} \frac{1}{1+r-\lambda_\ell} - \frac{1}{1+r} = \sum_{\ell=2}^{n} \frac{1}{1+r-\lambda_\ell}.$$

Similarly, we may define the *regularized Kemeny-based centrality measure* of the edge $e = \{i, j\}$

$$c_r(e) := K_r(G \setminus \{e\}) - K_r(G), \tag{5.1}$$

so that we have

$$c_r(e) = \sum_{\ell=2}^{n} \left(\frac{1}{1+r-\hat\lambda_\ell} - \frac{1}{1+r-\lambda_\ell} \right) = \sum_{\ell=2}^{n} \frac{\hat\lambda_\ell - \lambda_\ell}{(1+r-\hat\lambda_\ell)(1+r-\lambda_\ell)},$$

where $\hat\lambda_\ell$, $\ell = 1, \ldots, n$, are the eigenvalues of the matrix $D^{-1}\widehat{A}$, ordered in nonincreasing order, for $\widehat{A} = A + a_{ij}vv^T$ and $v = e_i - e_j$. Since $\hat\lambda_\ell \geq \lambda_\ell$ (compare the proof of Theorem 4.2), the above equation implies that $c_r(e) \geq 0$.

5.1. Interpretation of the regularized Kemeny constant. We can give an interpretation of the regularized Kemeny constant in terms of the random walk with teleportation that appears in the celebrated PageRank model [24]. We summarize briefly the model applied to the stochastic matrix $P = D^{-1}A$. Given a stochastic vector $w \in \mathbb{R}^n$ and a real number $\theta \in (0,1)$, the matrix $P^{(\theta)} = (1-\theta)P + \theta \mathbf{1} w^T$ is the transition matrix of the following Markov process: at each time instant, with probability $1-\theta$ we move from a node i to a node j according to the transition probabilities of the Markov chain associated to P, or with probability θ (informally called "teleportation probability") we move to a node chosen randomly and independently in the network according to the probability distribution given by the vector w.

If we set $\theta = \frac{r}{1+r}$, in view of the Brauer theorem [9], the eigenvalues of $P^{(\theta)}$ are $1, \frac{1}{1+r}\lambda_2, \frac{1}{1+r}\lambda_3, \ldots, \frac{1}{1+r}\lambda_n$, and thus

$$K(P^{(\frac{r}{1+r})}) = \sum_{i=2}^{n} \frac{1}{1 - \frac{1}{1+r}\lambda_i} = (1+r) \sum_{i=2}^{n} \frac{1}{1+r-\lambda_i} = (1+r) K_r(P).$$

Hence the regularized Kemeny constant of P is a constant multiple of the Kemeny constant of the Markov chain associated with the PageRank random walk with teleportation probability $\frac{r}{1+r}$. In this sense, the regularization can be interpreted as adding a small teleportation probability to the model.

5.2. Filtering procedure. If e is a cut-edge, then G is connected and $\widehat{G}$ is formed by two connected components; the matrix $\widehat{A}$ is reducible and has two eigenvalues equal to 1 so that $1 = \hat\lambda_1 = \hat\lambda_2 > \hat\lambda_3$. Thus, we may write

$$c_r(e) = r^{-1} - (1+r-\lambda_2)^{-1} + \sum_{\ell=3}^{n} ((1+r-\hat\lambda_\ell)^{-1} - (1+r-\lambda_\ell)^{-1}). \tag{5.2}$$

In this case, we have $\lim_{r \to 0} r c_r(e) = 1$ and it turns out that the regularized centrality measure of a cut-edge grows as r^{-1} when $r \to 0$. Observe also that the quantity $c_r(e)$ in (5.2) cannot exceed the value r^{-1}. In fact, we have the following result.

THEOREM 5.1. *Let G be a connected weighted undirected graph and e be a cut-edge of G. Then, for the regularized centrality measure of* (5.1) *we have $c_r(e) \leq r^{-1}$.*

Proof. Since e is a cut-edge, we may apply (5.2), which yields $c(e) - r^{-1} = \sum_{\ell=3}^{n}[(1+r-\hat{\lambda}_\ell)^{-1}-(1+r-\lambda_\ell)^{-1}]-(1+r-\lambda_2)^{-1}$, where λ_ℓ and $\hat{\lambda}_\ell$, $\ell=1,\dots,n$, are the eigenvalues of $D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ and of $D^{-\frac{1}{2}}\widehat{A}D^{-\frac{1}{2}}$, respectively, ordered in nonincreasing order. Thus we get

$$c_r(e) - r^{-1} = \sum_{\ell=2}^{n-1}[(1+r-\hat{\lambda}_{\ell+1})^{-1} - (1+r-\lambda_\ell)^{-1}] - 1/(1+r-\lambda_n). \tag{5.3}$$

Since $D^{-\frac{1}{2}}(\widehat{A}-A)D^{-\frac{1}{2}}$ is a positive semidefinite rank-one matrix, from the Cauchy interlacing property [22, section 47.4, Fact 1] we have $\hat{\lambda}_\ell \geq \lambda_\ell \geq \hat{\lambda}_{\ell+1}$ so that $(1+r-\hat{\lambda}_{\ell+1})^{-1} - (1+r-\lambda_\ell)^{-1} \leq 0$, which completes the proof in view of (5.3). ☐

Whenever the regularization parameter r is small enough, the additive term r^{-1} in (5.2) dominates, and the centrality measures of the cut-edges dominate the measures of the other edges. Moreover, cut-edges that connect two large components of a graph have roughly the same value r^{-1} of cut-edges that connect a single node to the remaining part of the graph.

Hence cut-edges have a huge value of the measure $c_r(e)$ irrespective of the mass of the graphs that they connect; a way to overcome this drawback can be obtained by modifying definition (5.1) as follows:

$$c_r^F(e) := \begin{cases} r^{-1} - c_r(e) & \text{if } e \text{ is a cut-edge,} \\ c_r(e) & \text{if } e \text{ is not a cut-edge,} \end{cases} \tag{5.4}$$

so that we still obtain nonnegative values in view of Theorem 5.1. We call $c_r^F(e)$ *filtered Kemeny-based centrality*. From (5.3) and (5.4), we deduce that $\lim_{r\to 0} c_r^F(e)$ is finite; moreover, if e is not a cut-edge, then $\lim_{r\to 0} c_r^F(e) = c(e)$.

In order to figure out if e is a cut-edge, one may apply the available computational techniques of [29], or, more simply, select those edges whose regularized score is of the order of r^{-1}. This can be achieved by means of a heuristic strategy by computing the unfiltered values $c_r(e)$ and selecting those edges e for which $c_r(e) > \frac{1}{2}r^{-1}$. Clearly this strategy works in the limit $r \to 0$, since $\lim_{r\to 0} c_r(e) = c(e)$ is finite whenever e is not a cut-edge, but it is not simple to determine a priori an explicit bound on r that ensures the success of the heuristic.

In Figure 5.1 we report the centrality measures c_r and c_r^F obtained by regularization and by filtered regularization, respectively, where thick blue edges denote high centrality values. This example shows the case where the removal of some edge splits the graph into two components. If $c_r(e) > \frac{1}{2}r^{-1}$, then e is considered a cut-edge. On the left, the centrality measure c_r with regularization parameter $r = 10^{-8}$ is computed. With this choice, the regularization parameter is approximately the square root of the machine precision. Observe that the regularized centrality values of cut-edges are of the order of r^{-1}. On the right, the filtering procedure has been applied. We may see that, in this case, only the edges that connect nonnegligible subgraphs have a higher value of the measure, but not of the order of r^{-1}, while the remaining disconnecting edges have an intermediate moderate value. At the moment, we have no theoretical explanation of this property that, on the other hand, appears clearly from the numerical experiments performed both on artificial data and on real world problems.

We see that, after the filtering procedure, the values $c_r^F(e)$ have comparable magnitudes across both cut-edges and non-cut-edges, and that their ordering matches remarkably well the intuitive notion of importance of an edge for the overall

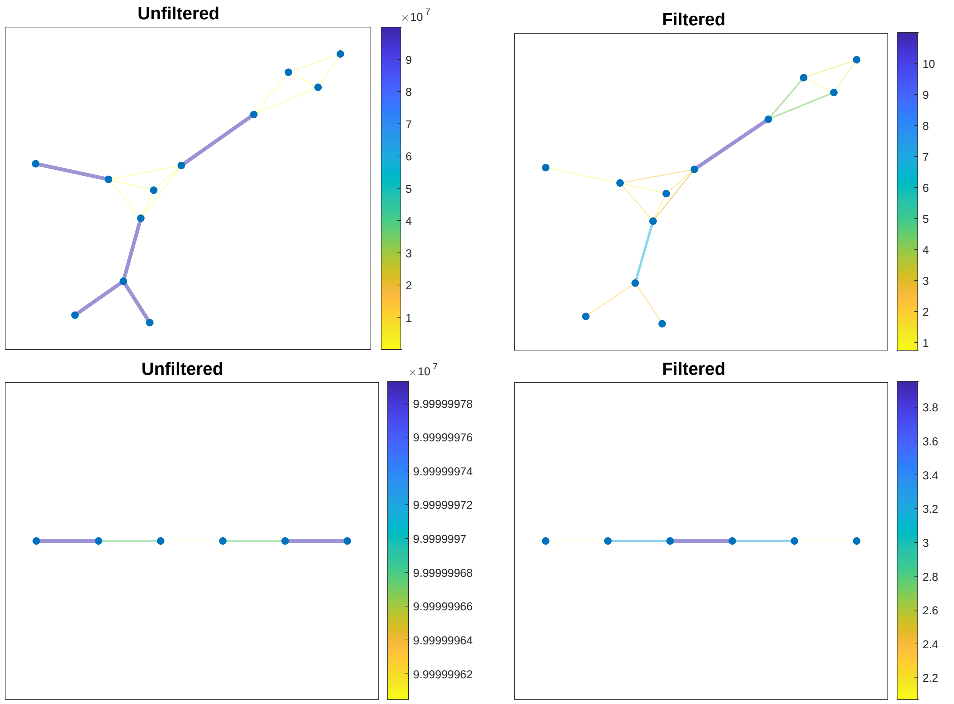

FIG. 5.1. *Centrality measures of two graphs having some cut-edges: On the left is the measure c_r computed with $r = 10^{-8}$; on the right is the filtered measure c_r^F. Blue and thick edges denote a higher centrality. Note the compressed color scale in the bottom-left figure: Since all edges are cut-edges, their unfiltered centrality measures are all very close to $r^{-1} = 10^8$. (Color available online.)*

connectivity of the graph. At the moment, we do not have theoretical results that relate $c_r^F(e_1)$ and $c_r^F(e_2)$ where e_1 is a cut-edge and e_2 is not.

6. Further theoretical results.

6.1. Extension to disconnected networks. If the graph G is disconnected, according to our earlier definitions, say, definition (2.2), we would get $K(G) = \infty$, since in this case the matrix P has at least two eigenvalues equal to 1. Therefore, one cannot apply the definition of the Kemeny-based centrality measure. However, we may extend this definition to disconnected networks by means of a continuity argument as follows.

Suppose that the graph G is disconnected; without loss of generality assume $P = \operatorname{diag}(P_1, P_2, \ldots, P_q)$, where P_ℓ, $\ell = 1, \ldots, q$, are irreducible stochastic matrices. Clearly, the matrix P has eigenvalues $\lambda_1 = \cdots = \lambda_q = 1$, and $\lambda_\ell \neq 1$ for $\ell = q+1, \ldots, n$. We consider the perturbed matrix $P^{(\theta)} := (1-\theta)P + \theta \mathbf{1} w^T$, again the random walk with teleportation associated with PageRank introduced in section 5.1. If w has positive entries, this matrix is stochastic and irreducible for any $0 < \theta \leq 1$, so that $P^{(\theta)}$ has only one eigenvalue $\lambda_1(\theta)$ equal to 1. Moreover, in view of the Brauer theorem [9], the remaining eigenvalues of $P^{(\theta)}$ are given by $\lambda_\ell(\theta) = (1-\theta)\lambda_\ell$, $\ell = 2, \ldots, n$. Therefore we have

$$K(P^{(\theta)}) = \sum_{\ell=2}^{n} \frac{1}{1-\lambda_\ell(\theta)} = \frac{q-1}{\theta} + \sum_{\ell=q+1}^{n} \frac{1}{1-(1-\theta)\lambda_\ell}.$$

Now consider the matrix $\widehat{P}^{(\theta)} = (1-\theta)\widehat{P} + \theta \mathbf{1} w^T$, where $\widehat{P}$ is obtained by removing the edge $\{i,j\}$. Assume that this edge belongs to the block P_s for some $1 \le s \le q$ and that it is not a cut-edge. That is, the block $\widehat{P}_s$ obtained after removing the edge is still irreducible. Denote by $\hat\lambda_\ell$, $\ell = 1,\dots,n$, the eigenvalues of $\widehat{P}$ so that $\hat\lambda_1 = \cdots = \hat\lambda_q = 1$, and $\hat\lambda_\ell \neq 1$ for $\ell = q+1,\dots,n$. By applying once again the Brauer theorem we find that $\widehat{P}^{(\theta)}$ has only one eigenvalue $\hat\lambda_1(\theta) = 1$, and the remaining eigenvalues are $\hat\lambda_\ell(\theta) = (1-\theta)\hat\lambda_\ell$ for $\ell = 2,\dots,n$. Therefore we have

$$K(\widehat{P}^{(\theta)}) = \sum_{\ell=2}^{n} \frac{1}{1-\hat\lambda_\ell(\theta)} = \frac{q-1}{\theta} + \sum_{\ell=q+1}^{n} \frac{1}{1-(1-\theta)\hat\lambda_\ell},$$

so that

$$K(\widehat{P}^{(\theta)}) - K(P^{(\theta)}) = \sum_{\ell=q+1}^{n} \left(\frac{1}{1-(1-\theta)\hat\lambda_\ell} - \frac{1}{1-(1-\theta)\lambda_\ell} \right),$$

whence

$$\lim_{\theta\to 0} \left(K(\widehat{P}^{(\theta)}) - K(P^{(\theta)}) \right) = \sum_{\ell=q+1}^{n} \left(\frac{1}{1-\hat\lambda_\ell} - \frac{1}{1-\lambda_\ell} \right).$$

Now recall that the removed entries $p_{i,j}$ and $p_{j,i}$ in $\widehat{P}$ belong to the block P_s so that the eigenvalues of $\widehat{P}$ different from the eigenvalues of P are those of the block P_s, except for the eigenvalue 1. Therefore we have

$$\lim_{\theta\to 0} \left(K(\widehat{P}^{(\theta)}) - K(P^{(\theta)}) \right) = K(\widehat{P}_s) - K(P_s).$$

From the above arguments we deduce the following result.

THEOREM 6.1. *Let G be a weighted undirected graph. Assume that G is not connected and that G_s is a connected component of G. Let e be an edge in G_s that is not a cut-edge in G_s. Denoting by $c_r(e)$ the regularized centrality measure of e in G and by $c(e)$ the centrality measure of e in G_s, we have*

$$\lim_{r\to 0} c_r(e) = c(e).$$

Therefore, the concept of the regularized centrality measure allows us to deal with disconnected graphs and avoids the search for connected components.

As an example of application, we consider the cases of two barbell-shaped graphs, together with their disjoint union. The pictures in the first row of Figure 6.1 represent the Kemeny-based centralities obtained by separately considering the two graphs, while the centralities obtained by considering the disjoint union of the graphs, where the adjacency matrix is reducible, are reported in the second row of Figure 6.1.

We can see from this representation that the centrality values of the disjoint union do not differ much from the union of the centralities of the two graphs. Moreover, observe that in the leftmost graph, where the barbell is formed by two loops, the edges in the loops have a high centrality. In fact, removing one of these edges almost disconnects the loop. Whereas, in the rightmost graph, where the loops are replaced by highly connected set of nodes, these edges have a low centrality. Indeed, their removal does not much alter the overall connectivity of the graph. On the other hand, removing one of the two edges connecting the two groups of nodes strongly reduces the connectivity between the two groups.

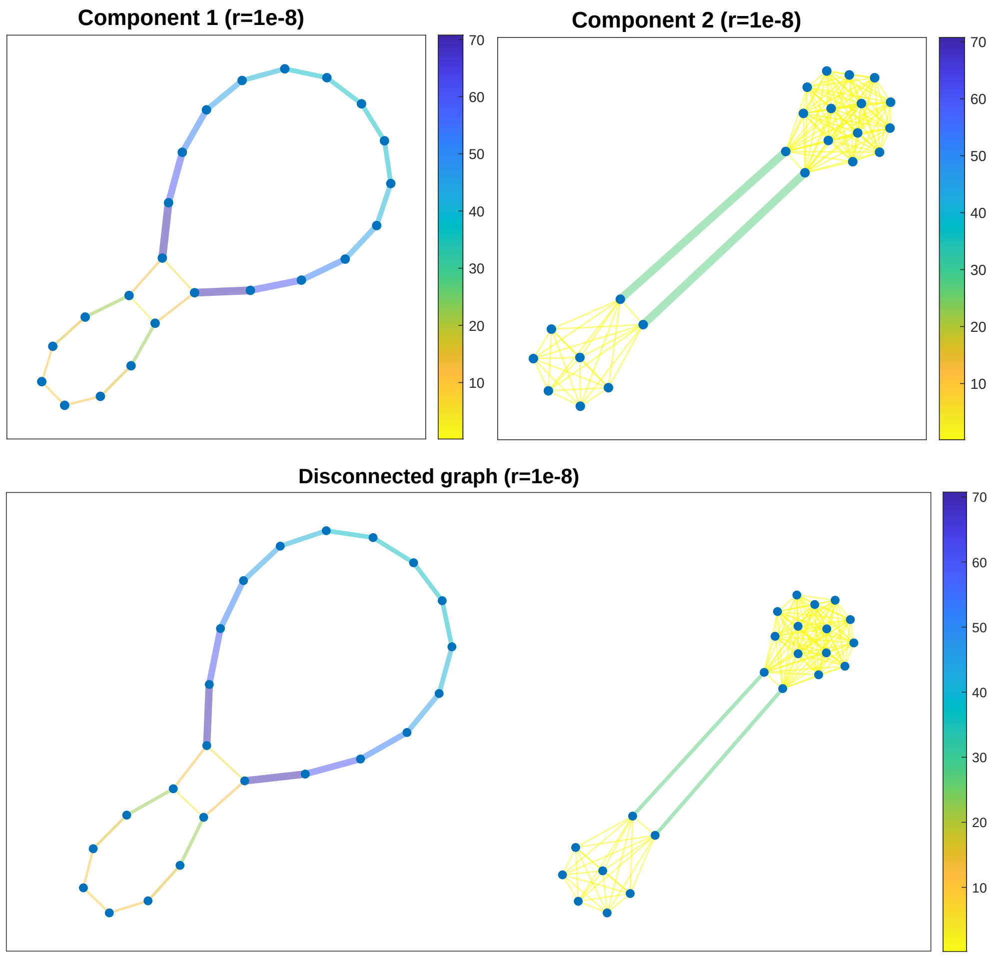

FIG. 6.1. *This figure shows the Kemeny-based regularized centrality measure applied separately to each connected component of a disconnected graph (first row) and to the union of the two components where the adjacency matrix is reducible (second row). Blue and thick edges denote higher Kemeny-based centrality. The values of the centralities of the disconnected graph do not differ much from those computed by separately applying the regularized measure to the two connected graph components. (Color available online.)*

6.2. Kemeny constant derivative. As above, let G be a connected weighted undirected graph, let $e=\{i,j\}\in E$ be an edge, and set $v=e_i-e_j$. It is natural to consider the function $f(t)=K(P(t))$, where $P(t)$ is the random walk transition matrix built starting from $A(t)=A+ta_{i,j}vv^T$ for $t\in[0,1]$. This function interpolates between the Kemeny constant of the original network $K(P)=f(0)$ and the one of the modified network $K(\widehat{P})=f(1)$. The centrality $c(e)$ can be written as the finite difference

$$c(e)=\frac{f(1)-f(0)}{1-0}.$$

The formula suggests studying the derivative $f'(t)=\lim_{\delta\to 0}\frac{f(t+\delta)-f(t)}{\delta}$, especially in $t=0$ and $t=1$, to better understand $c(e)$ and also possibly to replace it.

We have the following result.

THEOREM 6.2. *Under the assumptions of Theorem 4.2, let $f(t) = K(P(t))$ for $P(t) = D^{-1}A(t)$, where $A(t) = A + ta_{i,j}vv^T$, $D = \mathrm{diag}(d)$, $d = A(t)\mathbf{1} = A\mathbf{1}$, $t \in [0,1]$. Let λ_ℓ, $\ell = 1,\dots,n$, be the eigenvalues of $P(0)$, where $\lambda_1 = 1$. Then $0 \le f'(0) \le a_{i,j}(d_i^{-1} + d_j^{-1})\sum_{\ell=2}^n \frac{1}{(1-\lambda_\ell)^2}$.*

Proof. Denote by $\lambda_i(t)$ the eigenvalues of $C(t) := D^{-\frac{1}{2}}A(t)D^{-\frac{1}{2}} - \frac{1}{\|d\|_1}dd^T$. We have

$$\frac{1}{t}\big(K(P(t)) - K(P(0))\big) = \frac{1}{t}\sum_{\ell=2}^n \left(\frac{1}{1-\lambda_\ell(t)} - \frac{1}{1-\lambda_\ell}\right) = \sum_{\ell=2}^n \frac{(\lambda_\ell(t)-\lambda_\ell)/t}{(1-\lambda_\ell(t))(1-\lambda_\ell)}.$$

Since $\lambda_\ell \le \lambda_\ell(t) \le \lambda_\ell + ta_{i,j}(d_i^{-1} + d_j^{-1})$ (compare with the proof of Theorem 4.2), taking the limit for $t \to 0$ yields

$$f'(0) \le a_{i,j}(d_i^{-1} + d_j^{-1})\sum_{\ell=2}^n \frac{1}{(1-\lambda_\ell)^2}. \qquad \square$$

A similar inequality can be proved for $f'(1)$:

$$f'(1) \le a_{i,j}(d_i^{-1} + d_j^{-1})\sum_{\ell=2}^n \frac{1}{(1-\widehat{\lambda}_\ell)^2}.$$

Observe that the upper bound to $f'(0)$ given in the above theorem coincides with the value $a_{i,j}(d_i^{-1} + d_j^{-1})$ up to within a constant factor independent of i and j. This value depends on the degree of node i and of node j independently of the topology of the graph. The result suggests that in order to obtain a large change in the Kemeny constant one should focus on edges that not only have a large weight but also connect nodes with small degrees. Hence this argument supports the observation that the centrality measure $c(e)$ is effective in identifying weak ties in a network.

It is interesting to point out that in the paper [5] an expression of the derivative of the Kemeny constant is considered, to design a decomposition algorithm for the states of a Markov chain.

7. Computational aspects. In order to compute the regularized centrality $c_r(e)$ of an edge e, we may repeat the arguments of section 4 used to provide simple formulas for computing $c(e)$. In particular, the expression of $c(e)$ given in Theorem 4.3 still holds with W^{-1} and $\widehat{W}^{-1}$ replaced by $W_r^{-1} = (1+r)I - D^{-1/2}AD^{-1/2}$ and $\widehat{W}_r^{-1} = W_r^{-1} - a_{i,j}D^{-1/2}vv^TD^{-1/2}$, respectively. This way, equations (4.6) are still valid with S replaced by the positive definite matrix $S + rD$. That is, instead of inverting S directly, we may compute $F_r = (S+rD)^{-1}$ for a small positive value of the regularization parameter r. As pointed out in section 6.1, this regularization approach allows us to also treat the cases where the network is disconnected, so that the matrix S is singular, and the case where the removal of an edge disconnects the network. In that case, the corresponding θ in (4.6) coincides with 1.

The computation of the centrality measure of all the edges, with the regularization technique, is reported in Algorithm 7.1.

In this approach the amount of available RAM must be of the order of n^2 in order to store all the entries of S^{-1}. Indeed, large networks require a huge storage capacity.

A possible way to overcome the storage issues encountered in the case of large networks consists in exploiting the sparsity of the matrix A. In fact, since $T = (1+r)D - A$ is positive definite, there exists its Cholesky factorization $T = LL^T$.

Algorithm 7.1 Regularized Kemeny-based centrality of all the edges, where the number n of nodes is small enough so that n^2 entries can be stored in the RAM.

Input: The adjacency matrix A and a regularizing parameter $r > 0$
Output: The value $c_r(e)$ for any edge $e = \{i, j\}$
1: Compute $d = A\mathbf{1}$ and $\gamma = d^T \mathbf{1}$;
2: Set $D = \mathrm{diag}(d)$ and $S = (1+r)D - A + \frac{1}{\gamma} dd^T$;
3: Compute $F = S^{-1}$, $Q = F^T DF$;
4: **for all** edges $e = \{i, j\}$ **do**
5: compute $\alpha = a_{ij}(f_{ii} + f_{jj} - 2f_{ij})$, $\beta = a_{ij}(q_{ii} + q_{jj} - 2q_{ij})$, and $c_r(e) = \beta/(1-\alpha)$.
6: **end for**

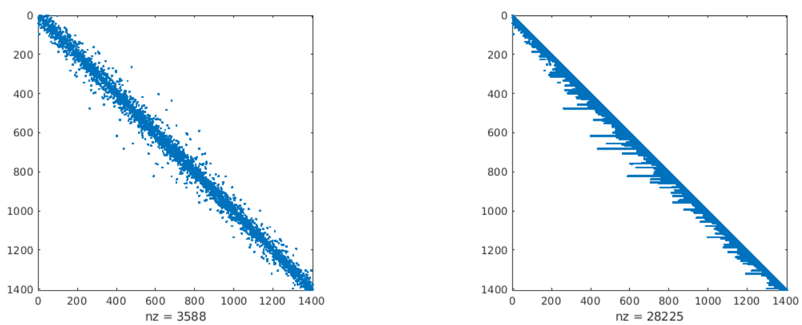

FIG. 7.1. *Plots (obtained with Matlab's* `spy`*) of the sparsity patterns of the adjacency matrix of the Pisa road network (left) and of its Cholesky factor (right). The number* `nz` *of nonzero entries is displayed as well.*

Moreover, the sparsity of T induces a sparsity structure in L so that the matrix L can be stored with a low memory space and the triangular systems having matrices L and L^T can be solved at a low cost. An example is given in Figure 7.1, where the structures of T and of L are displayed.

By applying the Sherman–Woodbury–Morrison identity to $S = T + \frac{1}{\|d\|_1} dd^T$ we may write

$$S^{-1} = T^{-1} - \frac{1}{\|d\|_1 + d^T z} zz^T, \quad z = T^{-1} d$$

so that for the vector x in (4.6) we have

$$x = w - \frac{z^T d}{\|d\|_1 + d^T z} z, \quad w = T^{-1} v.$$

Moreover, from the Cholesky factorization $T = LL^T$ we get

$$LL^T z = d, \quad LL^T w = v.$$

The above expressions can be used together with the first equation in (4.6) in order to compute $c_r(e)$.

Observe that from the computational point of view, one has to compute the Cholesky factorization once for all; this is cheaper than inverting a matrix. Moreover, two sparse triangular systems with matrix L and L^T must be solved once for all for computing z. Finally, for any edge $\{i, j\}$, two sparse triangular systems must be solved

Algorithm 7.2 Regularized and filtered Kemeny-based centrality of all the edges, relying on the Cholesky factorization.

Input: The adjacency matrix A and a regularizing parameter $r > 0$
Output: The value $c_r^F(e)$ for any edge $e = \{i, j\}$

```
Compute d = A1;
Set D = diag(d) and T = (1 + r)D − A;
Compute the Cholesky factorization T = LL^T;
Solve the linear systems Ly = d and L^T z = y;
Compute γ = d^T z + d^T 1;
for all edges e = {i, j} do
    set v = e_i − e_j
    solve the systems Ly = v and L^T w = y
    set δ = d^T w and x = w − (δ/γ) z
    compute α = a_ij (x_i − x_j), β = a_ij Σ_{ℓ=1}^n x_ℓ^2 d_ℓ, and c_r(e) = β/(1 − α);
    if c_r(e) > (1/2) r^{-1} then
        c_r^F(e) = r^{-1} − c_r(e)
    else
        c_r^F(e) = c_r(e)
    end if
end for
```

for computing w and $O(n)$ additional operations must be performed. Indeed, in this approach the cost is higher but this allows one to deal with large networks even if the amount of RAM storage is not sufficiently large. Algorithm 7.2 implements this approach, including regularization.

Note that after the precomputation steps the computation of the centrality of each edge is independent of the others; hence the main loop can be performed in parallel.

Besides the artificial tests shown in Figures 5.1 and 6.1, we have considered a real network composed of the roads in the city center of Pisa. To each edge $\{i, j\}$, corresponding to a road connecting i and j, we have assigned a connection strength $a_{ij} = \exp(-\mathcal{L}(i,j)/\mathcal{L}_{\max}) \in (0,1]$, where $\mathcal{L}(i,j)$ is the length of the edge, i.e., the Euclidean distance between points i and j, and $\mathcal{L}_{\max}$ is the maximum edge length in the network. We chose this formula for the weights to have a simple decreasing function of the length of the road to use in our tests. This is different from the approach taken in [14], where the authors work on the dual graph and use weights based on preexisting data on the real-world traffic to measure the connection strength. Here, we do not have traffic data to rely on, so we have to take a different automated approach.

This network is a planar undirected graph with 1794 nodes and 3240 edges; it includes many dead ends that are cut-edges, and various roads that, while not being cut-edges, are important bottlenecks for connectivity; among them are bridges on the Arno river and overpasses over the railroad line.

In our experiments, we have computed filtered and unfiltered Kemeny-based centralities with regularization of all the roads. We have compared these values to the following corresponding measures:

- The *Road-taking probability* in the PageRank model (with $\theta = 0.85$) is defined for an edge $\{i, j\}$ as $rt(\{i,j\}) = \pi_i R_{ij} + \pi_j R_{ji}$, where π is the PageRank vector and $R = \theta P + (1-\theta)\frac{1}{n}\mathbf{1}\mathbf{1}^T$ is the stochastic transition matrix of the PageRank model; this quantity corresponds to the long-term probability that a random surfer goes through that edge (in any direction).

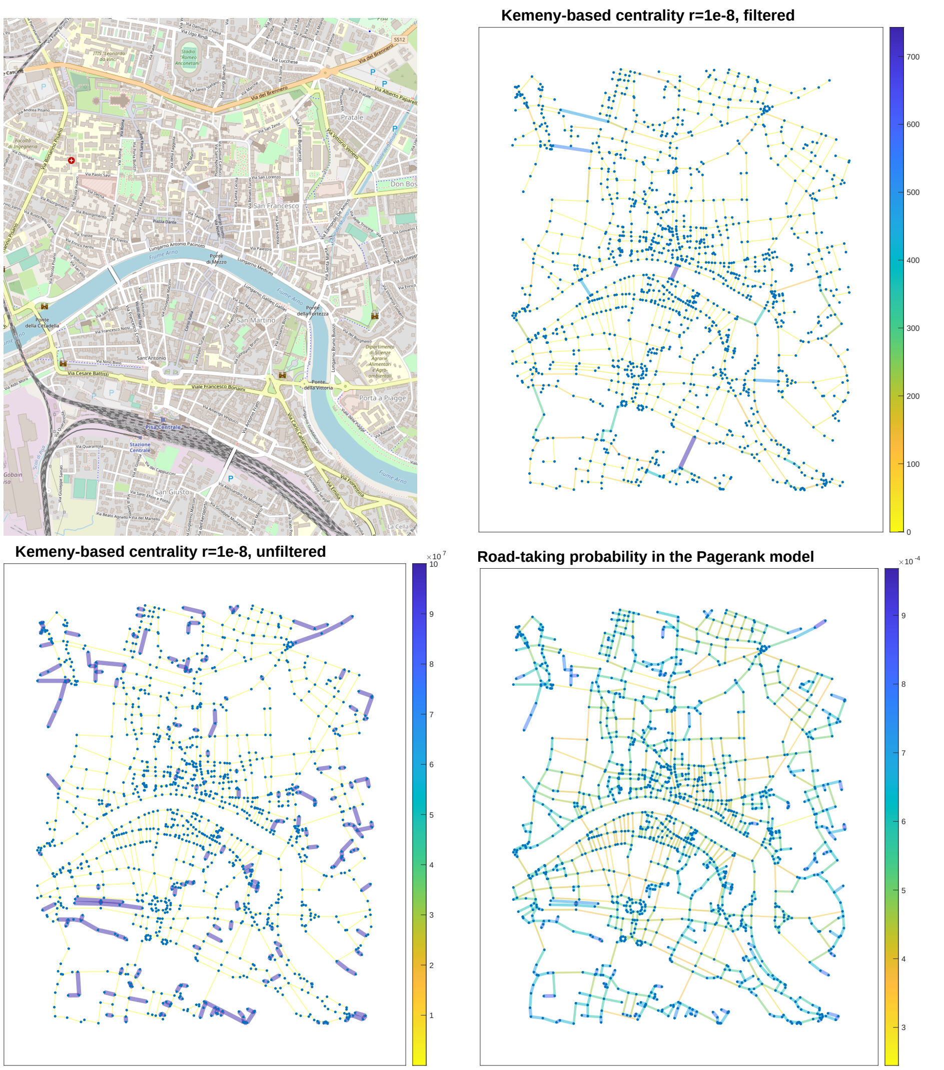

FIG. 7.2. *Comparison of several centrality measures on a map of the Pisa city center; part I.*

- *PageRank and Betweenness* on the dual graph are defined using the so-called *dual graph*, or *line graph*, of the network, i.e., a graph in which each road is a node, and two nodes are joined by an edge if the corresponding roads meet. This allows us to compute edge centrality measures with these two algorithms, which were designed to compute node importance.
- *Edge betweenness* (with $\mathcal{L}(i,j)$ as the distance), *Edge current-flow betweenness* [8] (with resistances a_{ij}), and *Edge load centrality*.

The PageRank based measures in the first two items are computed with the MATLAB `graph/centrality` command, while the Edge measures in the last item are computed using Python's NetworkX library [21]. Note that we have included in the comparison all edge centrality measures available in NetworkX, for completeness. We refer the reader to the book [16] and to the documentation of NetworkX for more details on these measures.

In Figure 7.2 we report the Pisa center road map (top left) and the results obtained by computing the regularized Kemeny-based centrality (filtered on top

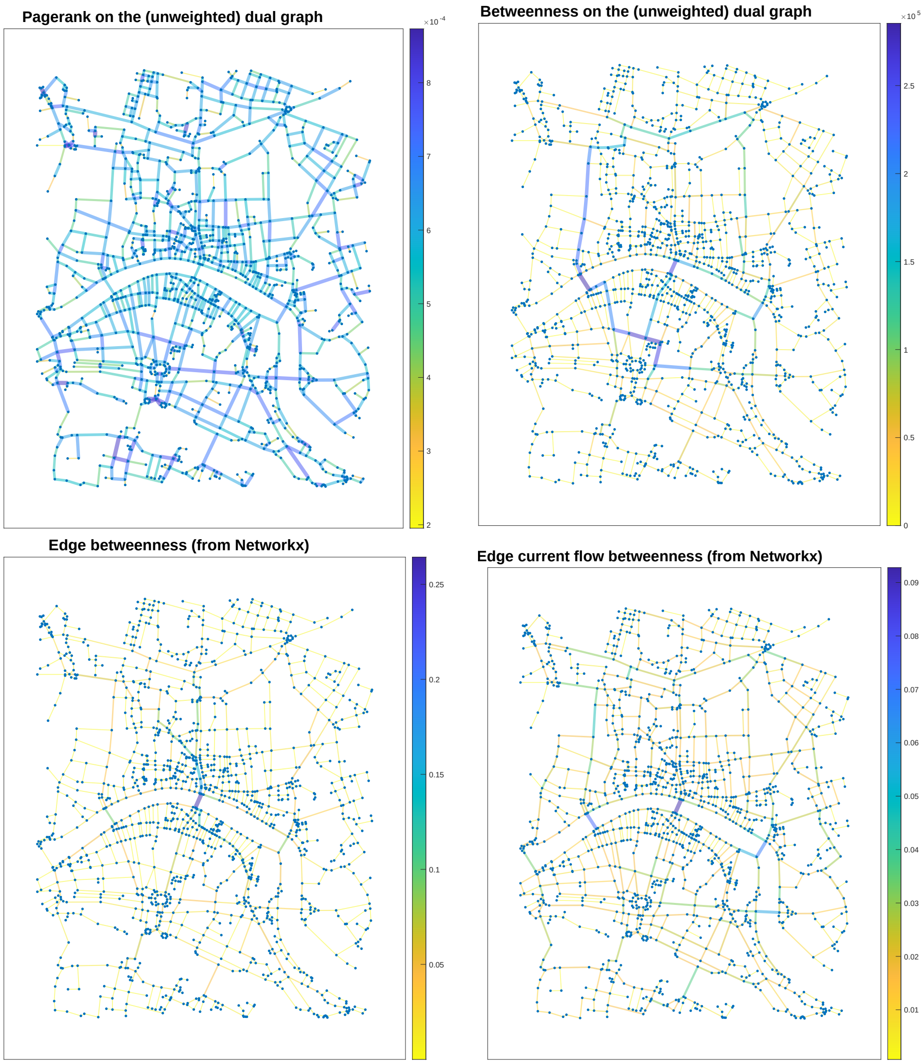

FIG. 7.3. *Comparison of several centrality measures on a map of the Pisa city center; part II.*

right; unfiltered on bottom left) and the Road-taking probability in the PageRank model.

We can see that the filtered version of the Kemeny-based centrality $c_r^F(e)$ (here computed with regularization parameter $r = 10^{-8}$) does an excellent job at highlighting important bottlenecks for connectivity (weak ties), such as bridges on the Arno river and overpasses over the railroad line. Moreover, dead ends that are cut-edges are not enhanced as important roads. In the unfiltered version of the measure, instead, cut-edges take a very high value and are essentially the only ones to be displayed in blue. This confirms that the filtering procedure is necessary to obtain sensible results. The Road-taking probability measure is not able to identify the important roads for connectivity.

In Figure 7.3 we display the results obtained with the other centrality measures. More specifically, PageRank on the unweighted dual graph (top right), Betweenness on the unweighted dual graph (top right), Edge betweenness (bottom left), and Edge

TABLE 7.1
CPU times for the various edge centrality algorithms on an Intel Core i5-1135G7 @ 2.40GHz laptop.

Algorithm	MATLAB	MATLAB R2021a with BGL [18]	Python 3.9.7 with NetworkX 2.4 [21]
Kemeny-based centrality	0.40 s		
Pagerank	0.001 s		0.19 s
Pagerank dual	0.001 s		0.1 s
Betweenness dual	0.04 s	0.64 s	6.21 s
Edge betweenness		0.42 s	6.51 s
Edge current-flow betweenness			7.3 s
Edge load centrality			6.8 s

current flow betweenness (bottom right). The results of Edge load centrality of [21] are not reported since they almost coincide with the ones of Betweenness on the unweighted dual graph and produce to an image that is indistinguishable to the eye.

From Figure 7.3, it turns out that PageRank and Betweenness, as well as Edge load centrality, fail to detect bottlenecks. Edge betweenness (with $\mathcal{L}(i,j)$ as the distance) and Edge current-flow betweenness partially succeed in highlighting bottleneck edges even though they do so at a much larger execution CPU time (see Table 7.1) if performed with Python's NetworkX library [21].

A detailed comparison in terms of CPU time is difficult due to the very nonuniform state of the codebases for the algorithms that we have used and the different overheads in the two (interpreted) languages that we have used; nevertheless, we report the observed timings in Table 7.1.

From Table 7.1, one can see that among the measures that better detect the bottlenecks, the filtered Kemeny-based centrality is the one that takes the smallest CPU time. Compared with the Python library NetworkX of [21], this measure is 12 times faster.

In theoretical terms, the time complexity of the Kemeny-based centrality is comparable with the cost of one matrix inversion, or the solution of n linear systems, which is $O(nm)$ when the Cholesky factor has $O(m)$ entries, as is the case for our road network examples. The complexity of Edge current-flow betweenness is similar, as it is also based on the pseudoinverse of the Laplacian, while Edge betweenness can be computed in $O(nm)$ as well for all networks. Measures computed using the dual graph have a similar cost because $m = O(n)$ for our road networks. PageRank-based measures are significantly cheaper, as they require the solution of only one linear system instead of n.

Another interesting observation is that, with the initial version of the measure $\widetilde{c}(e)$, 10.4% of the non-cut-edges in the Pisa map are Braess edges with $\widetilde{c}(e) < 0$. This shows that the Braess paradox has practical relevance and cannot be ignored as just an uncommon occurrence.

Finally, we describe the results of an experiment at a much larger scale. We have used the same methodology to compute the Kemeny-based centrality of a road network of the Tuscany region in Italy, a very large graph with 1.22M nodes and 1.56M edges. Despite the large size of the network, the Cholesky factor L is quite sparse, with only 3.36M nonzero entries, and it is computed in less than one second using MATLAB. A much more challenging computation is the computation of the centralities of each edge, each one of which requires solving two triangular linear systems with L and L^T. We have run this computation in parallel (using the MATLAB function `parfor`) on a machine with 12 physical cores with 3.4 GHz speed each (Intel Xeon CPU E5-2643) and MATLAB R2017a. The computation took 18 hours.

8. Conclusions. We have introduced a centrality measure for the edges of an undirected weighted graph based on the variation of the Kemeny constant. This measure has been modified in order to produce nonnegative results and avoid the presence of the Braess paradox, which would make the measure interpretation and its practical usage in applications more difficult. A regularization technique has been introduced for its computation; the technique allows one to detect cut-edges and to manage disconnected graphs. A technique of filtering has been introduced together with the filtered Kemeny-based centrality. From numerical experiments performed with artificial networks and with real road networks, it turned out that the filtered measure allows us to highlight weak ties, i.e., edges that connect large communities. All the versions of the Kemeny-based centrality measures can be expressed by means of the trace of suitable matrices, and their computation is ultimately reduced to the Cholesky factorization of a positive definite matrix, which is generally sparse. If the number of edges is huge, other techniques to estimate the trace of a matrix might be more appropriate, like the one proposed in [13] based on randomization. This is a subject of further research.

Our main interest here was to identify weak ties in road networks. But identifying weak ties has potential applications also in other fields. The seminal paper by Granovetter [20] considered social networks, for instance; and in the analysis of economical and biological networks this measure has the potential to reveal key connections that may go unnoticed under other metrics.

Code availability. MATLAB code to compute the centrality measure proposed in this paper, as well as MATLAB and Python code to reproduce the figures of this paper, is available at https://github.com/numpi/kemeny-based-centrality.

Acknowledgment. The authors wish to thank the reviewers for their careful reading and for many suggestions and remarks that substantially contributed to improving the presentation of the paper.

REFERENCES

[1] D. Altafini, D. Bini, V. Cutini, B. Meini, and F. Poloni, *Markov-chain based centralities and Space Syntax' Angular analysis: An initial overview and application*, in Proceedings of the 13th International Space Syntax Symposium, A. van Nes, and R. de Koning, eds., Western Norway University of Applied Sciences (HVL), Bergen, Norway, 2022.

[2] M. Benzi and P. Boito, *Matrix functions in network analysis*, GAMM-Mitt., 43 (2020), e202000012.

[3] M. Benzi, E. Estrada, and C. Klymko, *Ranking hubs and authorities using matrix functions*, Linear Algebra Appl., 438 (2013), pp. 2447–2474, https://doi.org/10.1016/j.laa.2012.10.022.

[4] M. Benzi and C. Klymko, *On the limiting behavior of parameter-dependent network centrality measures*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 686–706, https://doi.org/10.1137/130950550.

[5] J. Berkhout and B. Heidergott, *Analysis of Markov influence graphs*, Oper. Res., 67 (2019), pp. 892–904.

[6] R. Bhatia, *Matrix Analysis*, Grad. Texts in Math. 169, Springer-Verlag, New York, 1997, https://doi.org/10.1007/978-1-4612-0653-8.

[7] D. Braess, *Über ein paradoxon aus der verkehrsplanung*, Unternehmensforschung, 12 (1968), pp. 258–268, https://doi.org/10.1007/bf01918335.

[8] U. Brandes and D. Fleischer, *Centrality measures based on current flow*, in STACS 2005, 22nd Annual Symposium on Theoretical Aspects of Computer Science, V. Diekert and B. Durand, eds., Lecture Notes in Comput. Sci. 3404, Springer, Berlin, 2005, pp. 533–544, https://doi.org/10.1007/978-3-540-31856-9_44.

[9] A. Brauer, *Limits for the characteristic roots of a matrix*, VII, Duke Math. J., 25 (1958), pp. 583–590.

[10] J. Breen, S. Butler, N. Day, C. DeArmond, K. Lorenzen, H. Qian, and J. Riesen, *Computing Kemeny's constant for barbell-type graphs*, Electron. J. Linear Algebra, 35 (2019), pp. 583–598, https://doi.org/10.13001/1081-3810.4095.

[11] T. Bröhl and K. Lehnertz, *Centrality-based identification of important edges in complex networks*, Chaos, 29 (2019), 033115.

[12] T. Bröhl and K. Lehnertz, *A straightforward edge centrality concept derived from generalizing degree and strength*, Sci. Rep., 12 (2022), 4407.

[13] A. Cortinovis and D. Kressner, *On randomized trace estimates for indefinite matrices with an application to determinants*, Found. Comput. Math., 22 (2022), pp. 875–903, https://doi.org/10.1007/s10208-021-09525-9.

[14] E. Crisostomi, S. Kirkland, and R. Shorten, *A Google-like model of road network dynamics and its application to regulation and control*, Internat. J. Control, 84 (2011), pp. 633–651, https://doi.org/10.1080/00207179.2011.568005.

[15] O. De la Cruz Cabrera, M. Matar, and L. Reichel, *Edge importance in a network via line graphs and the matrix exponential*, Numer. Algorithms, 83 (2020), pp. 807–832, https://doi.org/10.1007/s11075-019-00704-y.

[16] E. Estrada, *The structure of complex networks: Theory and Applications*, Oxford Scholarship Online, 2013.

[17] M. Girvan and M. E. Newman, *Community structure in social and biological networks*, Proc. Natl. Acad. Sci. USA, 99 (2002), pp. 7821–7826.

[18] D. Gleich, *MatlabBGL. A Matlab graph library. Version* 4.0, 2008, https://www.cs.purdue.edu/homes/dgleich/packages/matlab_bgl/.

[19] G. H. Golub and C. F. Van Loan, *Matrix computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, MD, 1996.

[20] M. S. Granovetter, *The strength of weak ties*, Am. J. Sociol., 78 (1973), pp. 1360–1380.

[21] A. A. Hagberg, D. A. Schult, and P. J. Swart, *Exploring network structure, dynamics, and function using NetworkX*, in Proceedings of the 7th Python in Science Conference (SciPy2008), G. Varoquaux, T. Vaught, and J. Millman, eds., Pasadena, CA, 2008.

[22] L. Hogben, ed., *Handbook of Linear Algebra*, 2nd ed., Discrete Math. Appl., CRC Press, Boca Raton, FL, 2014.

[23] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*, reprint of 1960 original, Undergrad. Texts Math., Springer-Verlag, New York, Heidelberg, 1976.

[24] A. N. Langville and C. D. Meyer, *Google's PageRank and Beyond: The science of search engine rankings*, paperback ed. of the 2006 original, Princeton University Press, Princeton, NJ, 2012.

[25] M. Levene and G. Loizou, *Kemeny's constant and the random surfer*, Amer. Math. Monthly, 109 (2002), pp. 741–745, https://doi.org/10.2307/3072398.

[26] J. Lockhart, G. Minello, L. Rossi, S. Severini, and A. Torsello, *Edge centrality via the Holevo quantity*, in Structural, Syntactic, and Statistical Pattern Recognition, A. Robles-Kelly, M. Loog, B. Biggio, F. Escolano, and R. Wilson, eds., Springer, Cham, 2016, pp. 143–152.

[27] C. Mitchell, R. Agrawal, and J. Parker, *The effectiveness of edge centrality measures for anomaly detection*, in 2019 IEEE International Conference on Big Data (Big Data), IEEE, Washington, DC, 2019, pp. 5022–5027, https://doi.org/10.1109/BigData47090.2019.9006468.

[28] A. Schlote, E. Crisostomi, S. Kirkland, and R. Shorten, *Traffic modelling framework for electric vehicles*, Internat. J. Control, 85 (2012), pp. 880–897, https://doi.org/10.1080/00207179.2012.668716.

[29] R. E. Tarjan, *A note on finding the bridges of a graph*, Inform. Process. Lett., 2 (1973/74), pp. 160–161, https://doi.org/10.1016/0020-0190(74)90003-9.

[30] F. Tudisco and D. Higham, *Node and edge nonlinear eigenvector centrality for hypergraphs*, Commun. Phys., 4 (2021), 201.