Linear Algebra and its Applications 703 (2024) 137–162

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

On Kemeny's constant and stochastic complement [☆]

Dario Andrea Bini [a], Fabio Durastante [a,*], Sooyeong Kim [b], Beatrice Meini [a]

[a] Mathematics Department, University of Pisa, Largo Bruno Pontecorvo, 5, Pisa, 56127, PI, Italy
[b] Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, M3J 1P3, ON, Canada

ARTICLE INFO

Article history:
Received 20 December 2023
Received in revised form 30 August 2024
Accepted 1 September 2024
Available online 5 September 2024
Submitted by V. Mehrmann

MSC:
60J22
65C40
65F15

Keywords:
Markov chains
Kemeny's constant
Divide-and-conquer algorithm

ABSTRACT

Given a stochastic matrix P partitioned in four blocks P_{ij}, $i, j = 1, 2$, Kemeny's constant $\kappa(P)$ is expressed in terms of Kemeny's constants of the stochastic complements $P_1 = P_{11} + P_{12}(I - P_{22})^{-1}P_{21}$, and $P_2 = P_{22} + P_{21}(I - P_{11})^{-1}P_{12}$. Specific cases concerning periodic Markov chains and Kronecker products of stochastic matrices are investigated. Bounds to Kemeny's constant of perturbed matrices are given. Relying on these theoretical results, a divide-and-conquer algorithm for the efficient computation of Kemeny's constant of graphs is designed. Numerical experiments performed on real world problems show the high efficiency and reliability of this algorithm.

☆ This article was partially funded by the "INdAM – GNCS Project: Metodi basati su matrici e tensori strutturati per problemi di algebra lineare di grandi dimensioni" code CUP_E53C22001930001, and by the PRIN "Low-rank Structures and Numerical Methods in Matrix and Tensor Computations and their Application" code 20227PCCKZ. S. Kim is supported in part by funding from York-Fields Postdoctoral Fellowship grant and from the Natural Sciences and Engineering Research Council of Canada. The second and fourth authors are member of the INdAM GNCS group.

* Corresponding author.
E-mail addresses: dario.bini@unipi.it (D.A. Bini), fabio.durastante@unipi.it (F. Durastante), kimswim@yorku.ca (S. Kim), beatrice.meini@unipi.it (B. Meini).

https://doi.org/10.1016/j.laa.2024.09.001
0024-3795/

1. Introduction

Given an ergodic, finite, discrete-time, time-homogeneous Markov chain, Kemeny's constant is the expected time for the Markov chain to travel between randomly chosen states, where these states are sampled according to the stationary distribution. Originally defined in [22] as the expected time to reach a randomly-chosen state from a fixed starting state, this quantity is independent of the choice of initial state. Various intuitive explanations for the constancy have been provided in [3,14,27]. Moreover, Kemeny's constant has many applications to a variety of subjects, including the study of road traffic networks [1,11], disease spread [24,36], and many others.

In addition, Kemeny's constant has recently received significant attention within the graph theory community. Such constant finds special relevance in the study of random walks on graphs. In the context of random walks, Kemeny's constant is used as a measure of connectivity of the graph: the smaller the constant, the faster a random walker moves around the graph. As a graph invariant, much research has been dedicated to understanding how the structure of a graph influences Kemeny's constant. Extremal Kemeny's constant has been studied in [5,10,19]. The impact of edge addition/removal on this quantity has also been explored in [1,9,16,23,28,29]. Recently, work in [25] provides insights into the interplay between a graph and its complement regarding Kemeny's constant.

The pursuit of reducing the computational complexity of Kemeny's constant of a Markov chain with n states, which is $O(n^3)$, is interesting. Randomized approaches for the direct approximation of Kemeny's constant have been used in [30,35]. Additionally, exploring computational questions where partial information is available without computing from scratch has been investigated. Work in [1] investigated computation of Kemeny's constant for a graph obtained by removing an edge. Moreover, work from [6] provided an explicit formula for Kemeny's constant for graphs with bridges in terms of some quantities on the subgraphs resulting from the deletion of the bridges.

In this article, we furnish a formula for Kemeny's constant of a Markov chain by utilizing two Markov chains induced from the original, known as *censored (watched) Markov chains* (see [31]). Let P be the transition matrix of an irreducible Markov chain with state space $\{1, \ldots, n\}$. Let

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}, \tag{1}$$

where P_{11} and P_{22} are square of size $m \times m$ and $(n-m) \times (n-m)$, respectively. The transition matrices P_1 and P_2 of the censored Markov chains are given by

$$P_1 = P_{11} + P_{12}(I - P_{22})^{-1}P_{21}, \qquad P_2 = P_{22} + P_{21}(I - P_{11})^{-1}P_{12}. \tag{2}$$

In particular, we will provide expressions of the difference $\gamma = \kappa(P) - \kappa(P_1) - \kappa(P_2)$, given in terms of the stationary distribution vector of the Markov chain P.

The paper is organized as follows: In Section 2, we introduce some notation and review the definitions and properties of stochastic complements and Kemeny's constant. Section 3 concerns the main theoretical result (Theorem 3.3), where Kemeny's constant of a Markov chain is related to Kemeny's constants of censored Markov chains. Section 4 explores applications of the main result for various structured transition matrices and subsequently determines the minimum Kemeny's constant of a periodic Markov chain. In Section 5, we establish lower and upper bounds for the constant γ and bounds on Kemeny's constant of perturbed matrices are given. In Section 6, we introduce a divide-and-conquer algorithm to compute $\kappa(P)$ where P is a large and highly sparse matrix, and we present numerical experiments performed with real world matrices to show the effectiveness and reliability of this algorithm. Section 7 draws the conclusions.

2. Notation and preliminaries

A matrix $A = [a_{ij}] \in \mathbb{R}^{m\times n}$ is said to be *nonnegative (resp. positive)*, if $a_{ij} \geq 0$ ($a_{ij} > 0$) for any i, j, and we write $A \geq 0$ (resp. $A > 0$). We denote by $\mathbf{1}_n$ the all ones vector of size n. If n is clear from the context, we shall omit the subscript of $\mathbf{1}_n$. A matrix A is said to be *reducible* if there exists a permutation matrix Π such that $\Pi A \Pi^T$ is a block upper triangular matrix with square diagonal blocks. If A is not reducible, then we say that A is *irreducible*. A nonnegative matrix A is said to be *stochastic* if $A\mathbf{1} = \mathbf{1}$. For $A \in \mathbb{R}^{n\times n}$, we use $\rho(A)$ to denote the spectral radius of A.

A matrix A is a *non-singular (resp. singular)* M-matrix if $A = sI - B$ with $B \geq 0$ and $s > \rho(B)$ (resp. $s = \rho(B)$). It is known that if A is a non-singular M-matrix, then $a_{ii} > 0$, $a_{ij} \leq 0$ for $i \neq j$, and $A^{-1} \geq 0$.

Given a vector $\mathbf{v} = (v_i) \in \mathbb{R}^n$, we denote by $\|\mathbf{v}\|$ the 1-norm of $\mathbf{v}$, *i.e.*, $\|\mathbf{v}\| = \sum_{i=1}^n |v_i|$, and we use $\|\mathbf{v}\|_\infty$ to indicate the infinity norm of $\mathbf{v}$, that is, $\|\mathbf{v}\|_\infty = \max_i |v_i|$. For a matrix A, we denote by $\|A\|_\infty$ the norm induced by the infinity norm, that is, $\|A\|_\infty = \max_j \sum_{i=1}^n |a_{ij}|$.

2.1. Stochastic complement

Let P be an $n \times n$ irreducible stochastic matrix, with *stationary distribution vector* $\boldsymbol{\pi}$, *i.e.*, $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$ and $\boldsymbol{\pi}^T \mathbf{1} = 1$. Let P be partitioned into the 2×2 block matrix as in (1). Since P is irreducible, $I - P_{11}$ and $I - P_{22}$ are non-singular M-matrices. It is found in [31] that the matrices P_1 and P_2 in (2) are stochastic and irreducible. We call P_1 (resp. P_2) the *stochastic complement* of P_{11} (resp. P_{22}) in P. The matrix P_1 represents the transition matrix of the Markov chain obtained by censoring the states $\{m + 1, \ldots, n\}$, while P_2 represents the transition matrix of the Markov chain obtained by censoring the states $\{1, \ldots, m\}$. This suggests the name "censored Markov chain".

Let $\boldsymbol{\pi}$ be partitioned conformally with P so that $\boldsymbol{\pi}^T = \begin{bmatrix} \boldsymbol{\pi}_1^T & \boldsymbol{\pi}_2^T \end{bmatrix}$ where $\boldsymbol{\pi}_1 \in \mathbb{R}^m$ and $\boldsymbol{\pi}_2 \in \mathbb{R}^{n-m}$. Denoting by $\hat{\boldsymbol{\pi}}_i$ the stationary distribution vector of P_i, we have

$$\hat{\boldsymbol{\pi}}_i = \frac{1}{\|\boldsymbol{\pi}_i\|}\boldsymbol{\pi}_i, \quad i = 1, 2.$$

Moreover, setting $\alpha_i = \|\boldsymbol{\pi}_i\|$, we see that

$$\boldsymbol{\pi}^T = [\alpha_1 \hat{\boldsymbol{\pi}}_1^T, \alpha_2 \hat{\boldsymbol{\pi}}_2^T], \tag{3}$$

and the vector $\boldsymbol{\alpha}^T = [\alpha_1, \alpha_2]$ satisfies

$$\boldsymbol{\alpha}^T S = \boldsymbol{\alpha}^T, \quad \boldsymbol{\alpha}^T \mathbf{1} = 1, \quad S = \begin{bmatrix} \hat{\boldsymbol{\pi}}_1^T P_{11}\, \mathbf{1}_m & \hat{\boldsymbol{\pi}}_1^T P_{12}\, \mathbf{1}_{n-m} \\ \hat{\boldsymbol{\pi}}_2^T P_{21}\, \mathbf{1}_m & \hat{\boldsymbol{\pi}}_2^T P_{22}\, \mathbf{1}_{n-m} \end{bmatrix}. \tag{4}$$

That is, $\boldsymbol{\alpha}$ is the stationary distribution vector of the aggregated matrix S, which is stochastic and irreducible.

2.2. Kemeny's constant

Let P be the transition matrix of an ergodic, finite, discrete-time, time-homogeneous Markov chain. *Kemeny's constant* of P, denoted by $\kappa(P)$, is the expected number of time steps required by the Markov chain to go from a given state i to a random state j, sampled according to the stationary distribution $\boldsymbol{\pi}$. Kemeny's constant has an algebraic characterization, in terms of the trace of a suitable matrix [21,34], as stated in the following.

Lemma 2.1. *Let* $\mathbf{g}, \mathbf{h} \in \mathbb{R}^n$ *be vectors with* $\mathbf{h}^T\mathbf{g} = 1$, $\mathbf{h}^T\mathbf{1} \neq 0$, $\boldsymbol{\pi}^T\mathbf{g} \neq 0$. *Then,* $I - P + \mathbf{g}\mathbf{h}^T$ *is non-singular, and*

$$\kappa(P) = \operatorname{Tr}(Z) - \boldsymbol{\pi}^T Z\, \mathbf{1}, \tag{5}$$

where $Z = (I - P + \mathbf{g}\mathbf{h}^T)^{-1}$.

By choosing $\mathbf{g} = \mathbf{1}$, we find that

$$\kappa(P) = \operatorname{Tr}(Z) - 1 \quad \text{where} \quad Z = (I - P + \mathbf{1}\,\mathbf{h}^T)^{-1}. \tag{6}$$

Let $\lambda_1 = 1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of P. Since P is irreducible, λ_1 is simple. By the Brauer theorem [4], the eigenvalues $1 - \lambda_i$ of the matrix $I - P$ coincide with the eigenvalues of $I - P + \mathbf{1}\,\mathbf{h}^T$ except for the eigenvalue $1 - \lambda_1 = 0$ which is mapped into 1. This fact enables us to express Kemeny's constant in terms of the eigenvalues λ_i of P as

$$\kappa(P) = \sum_{i=2}^{n} \frac{1}{1 - \lambda_i}. \tag{7}$$

Recall that given two square stochastic matrices A, B of the same size, the products AB and BA are stochastic as well and have the same characteristic polynomial, see [17, Theorem 1.3.20]. So, their eigenvalues coincide and have the same algebraic multiplicities. Thus, it follows from (7) that

$$\kappa(AB) = \kappa(BA). \tag{8}$$

A slightly different situation is encountered for non-square matrices. Suppose that A and B are stochastic matrices of size $m \times n$ and $n \times m$, respectively, where $n > m$. Then AB and BA are square stochastic matrices of size $m \times m$ and $n \times n$, respectively. Moreover, their nonzero eigenvalues coincide, thus we deduce from (7) that

$$\kappa(BA) = \kappa(AB) + n - m. \tag{9}$$

2.3. Kemeny's constant and graphs

A graph is a pair $G = (V, E)$, where V is a finite set of *vertices* of cardinality $|V| = n$, and E is a set of pairs $\{i, j\}$ where $i, j \in V$, called *edges*. A graph is *weighted* if it is equipped with a function $a : E \to \mathbb{R}_+$ that associates a non-negative *weight* to each edge. For unweighted graphs, the weight of each edge is assumed to be equal to 1. The *adjacency matrix* of a graph on n vertices is the $n \times n$ matrix $A = [a_{ij}]$ such that a_{ij} is equal to the weight of edge $\{i, j\}$ if $\{i, j\} \in E$, and zero otherwise.

Given a graph G, we can associate the Markov chain having transition matrix $P = D^{-1}A$, where $D = \mathrm{diag}(\mathbf{d})$, $\mathbf{d} = A\,\mathbf{1}$, and A is the adjacency matrix of the graph. Here, we assume that $d_i \neq 0$ for any i, that is there are no isolated vertices. The matrix P describes a random walk on the graph G. We can then define Kemeny's constant of the graph as $\kappa(G) := \kappa(P)$.

A graph is said to be *bipartite* if there exist disjoint sets of vertices V_1 and V_2, of cardinality m_1 and m_2, respectively, such that $V = V_1 \cup V_2$, and the vertices in V_1 as well as the vertices in V_2 are connected by no edges. The adjacency matrix H of a bipartite graph can be partitioned into 4 blocks $H_{i,j}$, $i, j = 1, 2$ where H_{11} and H_{22} are the null square matrices of size m_1, and m_2, respectively. If the graph is unweighted complete bipartite then the blocks H_{12} and H_{21} of size $m_1 \times m_2$, $m_2 \times m_1$, respectively, have all the entries equal to 1.

3. Kemeny's constant and stochastic complement

In this section we relate Kemeny's constant of the transition matrix P in (1) to Kemeny's constants of the stochastic complements (2). We begin with introducing the following preliminary result.

Lemma 3.1. *Let P be the transition matrix given by (1), and let $Z = (I - P + \mathbf{1}\,\mathbf{h}^T)^{-1}$ where $\mathbf{h} \in \mathbb{R}^n$ with $\mathbf{h}^T\,\mathbf{1} = 1$. Partition $\mathbf{h}$ conformally with P as $\mathbf{h}^T = \begin{bmatrix} \mathbf{h}_1^T & \mathbf{h}_2^T \end{bmatrix}$. Then*

$$\mathrm{Tr}(Z) = \mathrm{Tr}(Z_1) + \mathrm{Tr}(Z_2) \tag{10}$$

where

$$Z_1 = (I - P_1 + \sigma_2^{-1}\mathbf{u}_1\mathbf{v}_1^T)^{-1}, \qquad Z_2 = (I - P_2 + \sigma_1^{-1}\mathbf{u}_2\mathbf{v}_2^T)^{-1}, \tag{11}$$

and

$$\begin{aligned} &\mathbf{u}_1 = \mathbf{1}_m + P_{12}(I - P_{22})^{-1}\,\mathbf{1}_{n-m}, \quad \mathbf{u}_2 = \mathbf{1}_{n-m} + P_{21}(I - P_{11})^{-1}\,\mathbf{1}_m, \\ &\mathbf{v}_1^T = \mathbf{h}_1^T + \mathbf{h}_2^T(I - P_{22})^{-1}P_{21}, \quad \mathbf{v}_2^T = \mathbf{h}_2^T + \mathbf{h}_1^T(I - P_{11})^{-1}P_{12}, \\ &\sigma_i = 1 + \mathbf{h}_i^T(I - P_{ii})^{-1}\,\mathbf{1} \quad \textit{for} \quad i = 1, 2. \end{aligned}$$

Proof. By partitioning Z conformally with P in (1), we find that

$$Z = \begin{bmatrix} I - P_{11} + \mathbf{1}_m\,\mathbf{h}_1^T & -P_{12} + \mathbf{1}_m\,\mathbf{h}_2^T \\ -P_{21} + \mathbf{1}_{n-m}\,\mathbf{h}_1^T & I - P_{22} + \mathbf{1}_{n-m}\,\mathbf{h}_2^T \end{bmatrix}^{-1}.$$

We shall omit the subscripts of all ones vectors. Then

$$Z = \begin{bmatrix} Z_1 & * \\ * & Z_2 \end{bmatrix},$$

where

$$\begin{aligned} Z_1 &= \left(I - P_{11} + \mathbf{1}\,\mathbf{h}_1^T - (P_{12} - \mathbf{1}\,\mathbf{h}_2^T)(I - P_{22} + \mathbf{1}\,\mathbf{h}_2^T)^{-1}(P_{21} - \mathbf{1}\,\mathbf{h}_1^T)\right)^{-1}, \\ Z_2 &= \left(I - P_{22} + \mathbf{1}\,\mathbf{h}_2^T - (P_{21} - \mathbf{1}\,\mathbf{h}_1^T)(I - P_{11} + \mathbf{1}\,\mathbf{h}_1^T)^{-1}(P_{12} - \mathbf{1}\,\mathbf{h}_2^T)\right)^{-1}, \end{aligned} \tag{12}$$

and $*$ denotes a generic entry. In particular, $\mathrm{Tr}(Z) = \mathrm{Tr}(Z_1) + \mathrm{Tr}(Z_2)$. By using the Sherman–Morrison formula, we find that

$$(I - P_{22} + \mathbf{1}\,\mathbf{h}_2^T)^{-1} = (I - P_{22})^{-1} - \sigma_2^{-1}(I - P_{22})^{-1}\,\mathbf{1}\,\mathbf{h}_2^T(I - P_{22})^{-1},$$

where $\sigma_2 = 1 + \mathbf{h}_2^T(I - P_{22})^{-1}\,\mathbf{1}$. Let $\mathbf{r}_1 = (I - P_{22})^{-1}\mathbf{1}$ and $\mathbf{r}_2^T = \mathbf{h}_2^T(I - P_{22})^{-1}$. Replacing this expression in the first formula in (12) yields

$$\begin{aligned} Z_1^{-1} = I - P_1 + \mathbf{1}\left(\left(1 - \mathbf{r}_2^T\mathbf{1} + \sigma_2^{-1}(\mathbf{r}_2^T\mathbf{1})(\mathbf{r}_2^T\mathbf{1})\right)\mathbf{h}_1^T + (1 - \sigma_2^{-1}\mathbf{r}_2^T\mathbf{1})\mathbf{r}_2^T P_{21}\right) \\ + \; P_{12}\mathbf{r}_1\left(\sigma_2^{-1}\mathbf{r}_2^T P_{21} + (1 - \sigma_2^{-1}\mathbf{r}_2^T\mathbf{1})\mathbf{h}_1^T\right). \end{aligned}$$

Since $\sigma_2 = 1 + \mathbf{r}_2^T\mathbf{1}$, we obtain the formula for Z_1 in (11). We proceed similarly for Z_2. $\square$

Observe that the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in Lemma 3.1 depend on the vector $\mathbf{h}$ that can be chosen arbitrarily under the condition $\mathbf{h}^T\,\mathbf{1} = 1$. In the next proposition we show that $\mathbf{h}$ can be chosen with $\mathbf{v}_1 = \hat{\boldsymbol{\pi}}_1$ and $\mathbf{v}_2 = \hat{\boldsymbol{\pi}}_2$. This choice will allow us to express Kemeny's constant of P in terms of Kemeny's constant of the stochastic complements P_1 and P_2, which will be shown in Theorem 3.3.

Proposition 3.2. *There exist vectors* $\mathbf{h}_1 \in \mathbb{R}^m$ *and* $\mathbf{h}_2 \in \mathbb{R}^{n-m}$ *such that*

$$\begin{aligned} \mathbf{h}_1^T + \mathbf{h}_2^T (I - P_{22})^{-1} P_{21} = \hat{\boldsymbol{\pi}}_1^T, \\ \mathbf{h}_2^T + \mathbf{h}_1^T (I - P_{11})^{-1} P_{12} = \hat{\boldsymbol{\pi}}_2^T, \end{aligned} \tag{13}$$

and $\mathbf{h}_1^T \mathbf{1}_m + \mathbf{h}_2^T \mathbf{1}_{n-m} = 1$.

Proof. Observe that (13) is equivalent to

$$\begin{aligned} &\mathbf{h}_1^T \left(I - (I - P_{11})^{-1} P_{12} (I - P_{22})^{-1} P_{21}\right) = \hat{\boldsymbol{\pi}}_1^T - \hat{\boldsymbol{\pi}}_2^T (I - P_{22})^{-1} P_{21}, \\ &\mathbf{h}_2^T = \hat{\boldsymbol{\pi}}_2^T - \mathbf{h}_1^T (I - P_{11})^{-1} P_{12}. \end{aligned} \tag{14}$$

Since the matrix $T = (I - P_{11})^{-1} P_{12} (I - P_{22})^{-1} P_{21}$ is stochastic, the vector $\mathbf{1}_m$ is orthogonal to the rows of $I - T$. On the other hand, since $(I - P_{22})^{-1} P_{21}$ is stochastic and $\hat{\boldsymbol{\pi}}_1^T \mathbf{1}_m = \hat{\boldsymbol{\pi}}_2^T \mathbf{1}_{n-m} = 1$, the vector $\mathbf{1}_m$ is orthogonal to the right-hand side of the first equation in (14). Hence, the system has a solution by the Rouché–Capelli Theorem [17, Section 0.2.4]. If $\mathbf{h}_1$ is a solution, then $\mathbf{h}_2$ can be recovered from the second equation in (14). By multiplying to the right by $\mathbf{1}_m$ the first equation in (13), we find that $\mathbf{h}_1^T \mathbf{1}_m + \mathbf{h}_2^T \mathbf{1}_{n-m} = 1$. □

Proposition 3.2, together with Equation (10) and Lemma 2.1, implies the following result.

Theorem 3.3. *Let* P *be a stochastic irreducible matrix, partitioned as in* (1)*, and denote by* P_1 *and* P_2 *the stochastic complements* (2)*. Then*

$$\kappa(P) = \kappa(P_1) + \kappa(P_2) + \gamma, \tag{15}$$

where

$$\gamma = \begin{bmatrix} \hat{\boldsymbol{\pi}}_1^T & -\hat{\boldsymbol{\pi}}_2^T \end{bmatrix} (I - P + \mathbf{u}\mathbf{v}^T)^{-1} \begin{bmatrix} \|\boldsymbol{\pi}_2\| \, \mathbf{1} \\ -\|\boldsymbol{\pi}_1\| \, \mathbf{1} \end{bmatrix}, \tag{16}$$

and $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ *are any vectors such that* $\mathbf{v}^T \mathbf{1} \neq 0$ *and* $\boldsymbol{\pi}^T \mathbf{u} \neq 0$.

Proof. From (6), $\kappa(P) = \mathrm{Tr}(Z) - 1$, where $Z = (I - P + \mathbf{1}\,\mathbf{h}^T)^{-1}$ and $\mathbf{h}$ is a vector such that $\mathbf{h}^T \mathbf{1} = 1$. Partition $\mathbf{h}$ as $\mathbf{h}^T = [\mathbf{h}_1^T, \mathbf{h}_2^T]$, where $\mathbf{h}_1$ and $\mathbf{h}_2$ satisfy (13). We shall maintain the same notation in Lemma 3.1. Then $\mathbf{v}_1^T = \hat{\boldsymbol{\pi}}_1^T$ and $\mathbf{v}_2^T = \hat{\boldsymbol{\pi}}_2^T$. Since $\hat{\boldsymbol{\pi}}_1^T (I - P_1) = 0$, the eigenvalues of $I - P_1 + \sigma_2^{-1} \mathbf{u}_1 \hat{\boldsymbol{\pi}}_1^T$ are the eigenvalues of $I - P_1$, except for the eigenvalue equal to zero, which is replaced by $\sigma_2^{-1} \hat{\boldsymbol{\pi}}_1^T \mathbf{u}_1$. Similarly, the eigenvalues of $I - P_2 + \sigma_1^{-1} \mathbf{u}_2 \hat{\boldsymbol{\pi}}_2^T$ are the eigenvalues of $I - P_2$, except for the eigenvalue equal to zero, which is replaced by $\sigma_1^{-1} \hat{\boldsymbol{\pi}}_2^T \mathbf{u}_2$. In view of (10), $\mathrm{Tr}(Z) = \mathrm{Tr}(Z_1) + \mathrm{Tr}(Z_2)$. From the expression of $\mathbf{u}_1$ in Lemma 3.1, we find that

$$\hat{\boldsymbol{\pi}}_1^T \mathbf{u}_1 = 1 + \hat{\boldsymbol{\pi}}_1^T P_{12}(I - P_{22})^{-1} \mathbf{1}_{n-m} = 1 + \frac{\boldsymbol{\pi}_2^T \mathbf{1}_{n-m}}{\|\boldsymbol{\pi}_1\|} = 1 + \frac{\|\boldsymbol{\pi}_2\|}{\|\boldsymbol{\pi}_1\|} = \frac{1}{\|\boldsymbol{\pi}_1\|}.$$

Hence, $\mathrm{Tr}(Z_1) = \kappa(P_1) + \sigma_2\|\boldsymbol{\pi}_1\|$. Similarly, $\hat{\boldsymbol{\pi}}_2^T \mathbf{u}_2 = \frac{1}{\|\boldsymbol{\pi}_2\|}$ and so $\mathrm{Tr}(Z_2) = \kappa(P_2) + \sigma_1\|\boldsymbol{\pi}_2\|$. Since $\mathrm{Tr}(Z) = \kappa(P_1)+\kappa(P_2)+\sigma_2\|\boldsymbol{\pi}_1\|+\sigma_1\|\boldsymbol{\pi}_2\|$, we have $\kappa(P) = \kappa(P_1)+\kappa(P_2)+\gamma$ where $\gamma = \sigma_2\|\boldsymbol{\pi}_1\| + \sigma_1\|\boldsymbol{\pi}_2\| - 1$.

Set $\mathbf{k}_i^T = \mathbf{h}_i^T(I - P_{ii})^{-1}$, $i = 1, 2$. Then

$$\gamma = \|\boldsymbol{\pi}_1\|\mathbf{k}_2^T \mathbf{1} + \|\boldsymbol{\pi}_2\|\mathbf{k}_1^T \mathbf{1}\,.$$

We find from (13) that $\mathbf{k}_1$ and $\mathbf{k}_2$ solve the linear system

$$\begin{bmatrix} \mathbf{k}_1^T & -\mathbf{k}_2^T \end{bmatrix} \begin{bmatrix} I - P_{11} & -P_{12} \\ -P_{21} & I - P_{22} \end{bmatrix} = \begin{bmatrix} \widehat{\boldsymbol{\pi}}_1^T & -\widehat{\boldsymbol{\pi}}_2^T \end{bmatrix}. \tag{17}$$

Since $(I - P)\mathbf{1} = 0$ and $\boldsymbol{\pi}^T(I - P) = 0$, this system is consistent. According to [21, Theorem 2], a solution of such a system is

$$\begin{bmatrix} \mathbf{k}_1^T & -\mathbf{k}_2^T \end{bmatrix} = \begin{bmatrix} \widehat{\boldsymbol{\pi}}_1^T & -\widehat{\boldsymbol{\pi}}_2^T \end{bmatrix} \widehat{Z}$$

where $\widehat{Z} = (I - P + \mathbf{u}\mathbf{v}^T)^{-1}$, and $\mathbf{u}$, $\mathbf{v} \in \mathbb{C}^n$ are such that $\mathbf{v}^T\mathbf{1} \neq 0$ and $\boldsymbol{\pi}^T\mathbf{u} \neq 0$. Therefore we arrive at (16). □

By choosing $\mathbf{u} = \mathbf{1}$ and $\mathbf{v} = \boldsymbol{\pi}$, we have

$$\begin{aligned}(I - P + \mathbf{1}\,\boldsymbol{\pi}^T)^{-1} \begin{bmatrix} \|\boldsymbol{\pi}_2\|\,\mathbf{1} \\ -\|\boldsymbol{\pi}_1\|\,\mathbf{1} \end{bmatrix} &= \left(I + \sum_{j=1}^{\infty} (P^j - \mathbf{1}\,\boldsymbol{\pi}^T) \right) \begin{bmatrix} \|\boldsymbol{\pi}_2\|\,\mathbf{1} \\ -\|\boldsymbol{\pi}_1\|\,\mathbf{1} \end{bmatrix} \\ &= \sum_{j=1}^{\infty} \left(P^j \begin{bmatrix} \|\boldsymbol{\pi}_2\|\,\mathbf{1} \\ -\|\boldsymbol{\pi}_1\|\,\mathbf{1} \end{bmatrix} \right)\end{aligned}$$

therefore

$$\gamma = \sum_{j=1}^{\infty} \left(\begin{bmatrix} \widehat{\boldsymbol{\pi}}_1^T & -\widehat{\boldsymbol{\pi}}_2^T \end{bmatrix} P^j \begin{bmatrix} \|\boldsymbol{\pi}_2\|\,\mathbf{1} \\ -\|\boldsymbol{\pi}_1\|\,\mathbf{1} \end{bmatrix} \right).$$

By partitioning P^j according to the partitioning of P, we have

$$P^j = \begin{bmatrix} P_{11}^{(j)} & P_{12}^{(j)} \\ P_{21}^{(j)} & P_{22}^{(j)} \end{bmatrix}.$$

Since $P^j\,\mathbf{1} = \mathbf{1}$, we find that

$$\begin{bmatrix} \widehat{\boldsymbol{\pi}}_1^T & -\widehat{\boldsymbol{\pi}}_2^T \end{bmatrix} P^j \begin{bmatrix} \|\boldsymbol{\pi}_2\|\,\mathbf{1} \\ -\|\boldsymbol{\pi}_1\|\,\mathbf{1} \end{bmatrix} = \widehat{\boldsymbol{\pi}}_1^T P_{11}^{(j)}\,\mathbf{1} + \widehat{\boldsymbol{\pi}}_2^T P_{22}^{(j)}\,\mathbf{1} - 1 = -\lambda_2^{(j)},$$

where $\lambda_2^{(j)}$ is the eigenvalue different from 1 of the 2×2 aggregated matrix

$$\begin{bmatrix} \widehat{\boldsymbol{\pi}}_1^T P_{11}^{(j)} \mathbf{1} & \widehat{\boldsymbol{\pi}}_1^T P_{12}^{(j)} \mathbf{1} \\ \widehat{\boldsymbol{\pi}}_2^T P_{21}^{(j)} \mathbf{1} & \widehat{\boldsymbol{\pi}}_2^T P_{22}^{(j)} \mathbf{1} \end{bmatrix}.$$

Since $\boldsymbol{\pi}^T(I - P + \mathbf{1}\,\boldsymbol{\pi}^T)^{-1} = \boldsymbol{\pi}^T$, $(I - P + \mathbf{1}\,\boldsymbol{\pi}^T)^{-1}\,\mathbf{1} = \mathbf{1}$, and $\boldsymbol{\pi}^T\,\mathbf{1} = 1$ we may conclude that

$$\gamma = \left(r\boldsymbol{\pi}^T + \begin{bmatrix} \widehat{\boldsymbol{\pi}}_1^T & -\widehat{\boldsymbol{\pi}}_2^T \end{bmatrix}\right)(I - P + \mathbf{1}\,\boldsymbol{\pi}^T)^{-1}\left(s\,\mathbf{1} + \begin{bmatrix} \|\boldsymbol{\pi}_2\|\,\mathbf{1} \\ -\|\boldsymbol{\pi}_1\|\,\mathbf{1} \end{bmatrix}\right) - rs,$$

for any scalars r and s. Other expressions for γ are stated by the following:

Proposition 3.4. *We have*

$$\gamma = \|\boldsymbol{\pi}_1\|\theta - \|\boldsymbol{\pi}_2\|, \tag{18}$$

where θ has one of the following equivalent expressions

$$\theta = \begin{bmatrix} \mathbf{0}^T & \hat{\boldsymbol{\pi}}_2^T \end{bmatrix}\left(I - P + \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix}\begin{bmatrix} \hat{\boldsymbol{\pi}}_1^T & \mathbf{0}^T \end{bmatrix}\right)^{-1}\begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix} \tag{19}$$

$$= \hat{\boldsymbol{\pi}}_2^T\left(I + (I - P_{22})^{-1}P_{21}(I - P_1 + \mathbf{1}\,\hat{\boldsymbol{\pi}}_1^T)^{-1}P_{12}\right)(I - P_{22})^{-1}\,\mathbf{1} \tag{20}$$

$$= \hat{\boldsymbol{\pi}}_2^T\left(I - P_{22} - P_{21}(I - P_{11} + \mathbf{1}\,\hat{\boldsymbol{\pi}}_1^T)^{-1}P_{12}\right)^{-1}\mathbf{1} \tag{21}$$

$$= \hat{\boldsymbol{\pi}}_2^T\left(I - P_2 + \frac{1}{1 + \hat{\boldsymbol{\pi}}_1^T S\,\mathbf{1}}P_{21}S\,\mathbf{1}\,\hat{\boldsymbol{\pi}}_1^T S P_{12}\right)^{-1}\mathbf{1}, \quad S = (I - P_{11})^{-1}. \tag{22}$$

Proof. From (16), by choosing $\mathbf{u}^T = \begin{bmatrix} \mathbf{1}^T & \mathbf{0}^T \end{bmatrix}$ and $\mathbf{v}^T = \begin{bmatrix} \boldsymbol{\pi}_1^T & \mathbf{0}^T \end{bmatrix}$, and by observing that $\widehat{Z}\mathbf{u} = \mathbf{1}$ and $\mathbf{v}^T\widehat{Z} = \boldsymbol{\pi}^T$, where $\widehat{Z} = (I - P + \mathbf{u}\mathbf{v}^T)^{-1}$, we arrive at (18), with θ given in (19). The remaining expressions are obtained by applying formal manipulations to (19) and properties of Schur complements. □

4. Application of Theorem 3.3

Here we apply the main result of the previous section to structured stochastic matrices.

4.1. Periodic Markov chains

Let P be the transition matrix of an ergodic Markov chain with n states. The *period* p_i of state i is the greatest common divisor of all natural numbers m such that $(P^m)_{i,i} > 0$. A Markov chain is called *periodic* if $p_i \geq 2$ for all $1 \leq i \leq n$, and it is called *aperiodic* otherwise. The *period* of a periodic Markov chain is the greatest common divisor of periods of all states. It is well-known that if the Markov chain is periodic, then P is permutationally similar to a block-cyclic matrix (see [8]).

Given a periodic Markov chain with n states and period d, we may assume that the transition matrix P is given by

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \left[\begin{array}{c|cccc} 0 & 0 & \dots & 0 & A_d \\ \hline A_1 & 0 & \ddots & \ddots & 0 \\ 0 & A_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & A_{d-1} & 0 \end{array}\right] \tag{23}$$

where all block diagonal matrices are square and each of A_i's is a rectangular stochastic matrix. Let n_i be the size of i^{th} block diagonal matrix of P for $i = 1, \dots, d$. A set of the states corresponding to a block diagonal matrix is called a *cyclic class*. The *cyclicity index* is the number of cyclic classes. One can find that each of

$$A_d A_{d-1} \cdots A_1, \quad A_1 A_d \cdots A_2, \quad \dots \quad , \quad A_{d-1} \cdots A_1 A_d \tag{24}$$

is a square stochastic matrix and its corresponding Markov chain is aperiodic. It can be seen that

$$(I - P_{22})^{-1} = \left[\begin{array}{c|c|c|c|c} I & 0 & & 0 & 0 \\ & & & \vdots & \vdots \\ A_2 & I & & & \\ A_3 A_2 & A_3 & \cdots & 0 & 0 \\ \vdots & \vdots & & I & 0 \\ A_{d-1} \cdots A_3 A_2 & A_{d-1} \cdots A_4 A_3 & & A_{d-1} & I \end{array}\right]. \tag{25}$$

It is immediate to find that the stochastic complements P_1 and P_2 of P_{11} and P_{22} are, respectively,

$$P_1 = A_d A_{d-1} \cdots A_1, \quad P_2 = \begin{bmatrix} 0 & \dots & \dots & 0 & A_1 A_d \\ A_2 & 0 & \ddots & \ddots & 0 \\ 0 & A_3 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & A_{d-1} & 0 \end{bmatrix}. \tag{26}$$

We note that the censored Markov chain for P_2 is periodic.

Proposition 4.1. *Let P be the transition matrix of a periodic Markov chain. Suppose that P is of form* (23)*. Then*

$$\kappa(P) = \kappa(P_1) + \kappa(P_2) + \frac{1}{2}$$

where P_1 and P_2 are given in (26).

Proof. It suffices to show that γ in (15) is $\frac{1}{2}$. In order to find γ, we shall find the value of θ introduced in Equation (18). From (19), we may look at θ as the sum of two terms θ_1 and θ_2, where

$$\begin{aligned}
\theta_1 &= \hat{\boldsymbol{\pi}}_2^T (I - P_{22})^{-1} \mathbf{1},\\
\theta_2 &= \hat{\boldsymbol{\pi}}_2^T (I - P_{22})^{-1} P_{21} (I - P_1 + \mathbf{1}\, \hat{\boldsymbol{\pi}}_1^T)^{-1} P_{12} (I - P_{22})^{-1} \mathbf{1}\,.
\end{aligned}$$

We first find expressions of $\hat{\boldsymbol{\pi}}_2^T$ and $(I - P_{22})^{-1}\mathbf{1}$. Let

$$\boldsymbol{\pi}^T = [\mathbf{s}_1^T, \mathbf{s}_2^T, \ldots, \mathbf{s}_d^T]$$

be conformally partitioned with P and be the stationary distribution vector. From the condition $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$, we have $\mathbf{s}_i^T = \mathbf{s}_{i+1}^T A_i$, $i = 1, \ldots, d-1$, and $\mathbf{s}_d^T = \mathbf{s}_1^T A_d$. Then

$$\mathbf{s}_2^T = \mathbf{s}_1^T A_d A_{d-1} \cdots A_2, \quad \mathbf{s}_3^T = \mathbf{s}_1^T A_d A_{d-1} \cdots A_3, \quad \ldots \quad , \quad \mathbf{s}_{d-1}^T = \mathbf{s}_1^T A_d A_{d-1}.$$

Moreover, since $A_i \mathbf{1} = \mathbf{1}$, it follows that $\mathbf{s}_i^T \mathbf{1} = \mathbf{s}_{i+1}^T \mathbf{1}$ and $\mathbf{s}_d^T \mathbf{1} = \mathbf{s}_1^T \mathbf{1}$. Hence, $\| \mathbf{s}_j \| = \frac{1}{d}$ for $1 \le j \le d$. Note that $\boldsymbol{\pi}_1 = \mathbf{s}_1$. Therefore,

$$\boldsymbol{\pi}^T = \frac{1}{d} \begin{bmatrix} \hat{\boldsymbol{\pi}}_1^T & \hat{\boldsymbol{\pi}}_1^T (A_d A_{d-1} \cdots A_2) & \hat{\boldsymbol{\pi}}_1^T (A_d A_{d-1} \cdots A_3) & \cdots & \hat{\boldsymbol{\pi}}_1^T A_d \end{bmatrix},$$

whence

$$\hat{\boldsymbol{\pi}}_2^T = \frac{1}{d-1} \begin{bmatrix} \hat{\boldsymbol{\pi}}_1^T (A_d A_{d-1} \cdots A_2) & \hat{\boldsymbol{\pi}}_1^T (A_d A_{d-1} \cdots A_3) & \cdots & \hat{\boldsymbol{\pi}}_1^T A_d \end{bmatrix}. \tag{27}$$

Observing $(I - P_{22})^{-1}$ in (25), we have

$$\begin{aligned}
&(I - P_{22})^{-1} \mathbf{1} = \begin{bmatrix} \mathbf{1}^T & 2\,\mathbf{1}^T & \cdots & (d-1)\,\mathbf{1}^T \end{bmatrix}^T,\\
&P_{12} (I - P_{22})^{-1} \mathbf{1} = (d-1)\,\mathbf{1},\\
&(I - P_{22})^{-1} P_{21} \mathbf{1} = \mathbf{1}\,.
\end{aligned} \tag{28}$$

By using the expression of $\hat{\boldsymbol{\pi}}_2$ in (27) and the expression of $(I - P_{22})^{-1}\mathbf{1}$ in (28) together with the property $A_i \mathbf{1} = \mathbf{1}$, we find that

$$\theta_1 = \frac{1}{d-1} \sum_{i=1}^{d-1} i \hat{\boldsymbol{\pi}}_1^T \mathbf{1} = \frac{1}{2} d.$$

Note that since $(I - P_1 + \mathbf{1}\,\hat{\boldsymbol{\pi}}_1^T)\,\mathbf{1} = \mathbf{1}$, we have $(I - P_1 + \mathbf{1}\,\hat{\boldsymbol{\pi}}_1^T)^{-1}\,\mathbf{1} = \mathbf{1}$. Concerning θ_2, we find from (28) that

$$\theta_2 = \hat{\boldsymbol{\pi}}_2^T (I - P_{22})^{-1} P_{21} (I - P_1 + \mathbf{1}\,\hat{\boldsymbol{\pi}}_1^T)^{-1} P_{12} (I - P_{22})^{-1}\,\mathbf{1} = d - 1.$$

Thus, $\theta = \theta_1 + \theta_2 = \frac{3}{2}d - 1$ and therefore $\gamma = \|\boldsymbol{\pi}_1\|\theta - \|\boldsymbol{\pi}_2\| = \frac{1}{2}$. □

Theorem 4.2. *Let P be the transition matrix of a periodic Markov chain. Suppose that P is of form (23). Then*

$$\kappa(P) = d\kappa(A_d A_{d-1} \cdots A_1) + n - dn_1 + \frac{d-1}{2}. \tag{29}$$

Moreover, if $n_1 = \cdots = n_d$ then

$$\kappa(P) = d\kappa(A_d A_{d-1} \cdots A_1) + \frac{d-1}{2}. \tag{30}$$

Proof. Consider P_2 in (26) with $d \geq 3$. Note that the Markov chain corresponding to P_2 is periodic. Applying Proposition 4.1, we find that

$$\kappa(P_2) = \kappa(A_1 A_d A_{d-1} \cdots A_2) + \kappa(P_3) + \frac{1}{2},$$

where

$$P_3 = \left[\begin{array}{c|ccc} 0 & \dots & 0 & A_2 A_1 A_d \\ \hline A_3 & \ddots & \ddots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & & A_{d-1} & 0 \end{array}\right].$$

Hence, $\kappa(P) = \kappa(A_d A_{d-1} \cdots A_1) + \kappa(A_1 A_d A_{d-1} \cdots A_2) + \kappa(P_3) + 1$. If $d \geq 4$, one can apply the proposition to P_3 with the partition. In this manner, recursively applying the proposition, we obtain

$$\begin{aligned} \kappa(P) = {} & \kappa(A_d A_{d-1} \cdots A_1) + \kappa(A_1 A_d A_{d-1} \cdots A_2) + \cdots \\ & + \kappa(A_{d-1} A_{d-2} \cdots A_1 A_d) + \frac{d-1}{2}. \end{aligned}$$

Recall that A_i is of size $n_{i+1} \times n_i$ for $1 \leq i \leq d-1$ and A_d is of size $n_1 \times n_d$. Note that $A_{i-1} A_{k-2} \cdots A_1 A_d A_{d-1} \cdots A_i$ is of size $n_i \times n_i$ for $1 \leq i \leq d$. From (9), we obtain

$$\begin{aligned} \kappa(A_{k-1} A_{k-2} \cdots A_1 A_d A_{d-1} \cdots A_k) = {} & \kappa(A_{k-2} A_{k-3} \cdots A_1 A_d A_{d-1} \cdots A_{k-1}) \\ & + n_k - n_{k-1} \end{aligned}$$

for $2 \leq k \leq d$. It follows that

$$\begin{aligned}&\kappa(A_dA_{d-1}\cdots A_1)+\cdots+\kappa(A_{d-2}\cdots A_1A_dA_{d-1})+\kappa(A_{d-1}\cdots A_1A_d)\\&\quad= d\kappa(A_dA_{d-1}\cdots A_1)+(d-1)(n_2-n_1)+\cdots\\&\quad+ 2(n_{d-1}-n_{d-2})+n_d-n_{d-1}= d\kappa(A_dA_{d-1}\cdots A_1)+n-dn_1. \quad\square\end{aligned}$$

Remark 4.3. The characteristic polynomials of the matrices P in (23) and $P_1 = A_dA_{d-1}\cdots A_1$ are related by the equation $\det(\lambda I - P) = \lambda^\ell \det(\lambda^d I - P_1)$, where $\ell = n - dn_1$. Therefore, from the computational point of view, if the eigenvalues of P_1 are explicitly known, then also the eigenvalues of P are explicitly known and we can recover $\kappa(P)$ from (7). On the other hand, if the eigenvalues of P_1 are not known, but the value of $\kappa(P_1)$ is available, then Kemeny's constant of P can be directly obtained from (29).

Example 4.4. A collection of random walks on undirected graphs is one of the most accessible families of Markov chains. If a random walk is periodic, then the underlying graph is necessarily bipartite. Let P be an irreducible stochastic matrix with the structure

$$P = \begin{bmatrix} 0 & P_{12} \\ P_{21} & 0 \end{bmatrix}$$

where P_{12} and P_{21} have size $n_1 \times n_2$ and $n_2 \times n_1$, respectively. Then

$$\kappa(P) = 2\kappa(P_1) - n_1 + n_2 + \frac{1}{2}, \tag{31}$$

where $P_1 = P_{12}P_{21}$.

Example 4.5. Consider the transition matrix P in (23). Suppose that for $i = 1, \ldots, d$, A_i is the matrix with all entries equal to $\frac{1}{n_i}$. Then $A_dA_{d-1}\cdots A_1$ is the $n_1 \times n_1$ matrix with all entries equal to $\frac{1}{n_1}$. Since $\kappa(A_dA_{d-1}\cdots A_1) = n_1 - 1$, we have

$$\kappa(P) = n - \frac{d+1}{2}.$$

Now we provide a lower bound on Kemeny's constant of a periodic Markov chain.

Corollary 4.6. *Let P be the transition matrix of a periodic Markov chain. Suppose that $P = [p_{xy}]$ is of form (23) with $n_1 \leq n_i$ for $1 \leq i \leq d$. Then*

$$\kappa(P) \geq n - \frac{dn_1 + 1}{2},$$

with equality when p_{xy} is given as in (32).

Proof. From (29), $\kappa(P) = d\kappa(A_d A_{d-1} \cdots A_1) + n - dn_1 + \frac{d-1}{2}$. It is known in [26, Remark 2.14] that given an $m \times m$ irreducible stochastic matrix P, $\kappa(P) \geq \frac{m-1}{2}$ with equality if and only if P is the adjacency matrix of a directed m-cycle. Hence, it is enough to find a periodic Markov chain such that $A_d A_{d-1} \cdots A_1$ is the adjacency matrix of a directed n_1-cycle.

Note that $n_1 \leq n_i$ for $1 \leq i \leq d$. We may suppose that for $j = 1, \ldots, d$, the cyclic class corresponding to j^{th} block diagonal matrix of $P = [p_{xy}]$ is partitioned into n_1 subsets $C_1^j, \ldots, C_{n_1}^j$. Let p_{xy} be given as follows:

$$p_{xy} = \begin{cases} \frac{1}{|C_\ell^d|}, & \text{if } x \in C_\ell^1 \text{ and } y \in C_\ell^d \text{ for } 1 \leq \ell \leq n_1; \\ \frac{1}{|C_\ell^{j-1}|}, & \text{if } x \in C_\ell^j \text{ and } y \in C_\ell^{j-1} \text{ for } 1 \leq \ell \leq n_1 \text{ and } 3 \leq j \leq d; \\ 1, & \text{if } x \in C_\ell^2 \text{ and } y \in C_{\ell-1}^1 \text{ for } 2 \leq \ell \leq n_1, \text{ or } x \in C_1^2 \text{ and } y \in C_{n_1}^1; \\ 0, & \text{otherwise.} \end{cases} \tag{32}$$

It can be seen that $A_d A_{d-1} \cdots A_1$ is the adjacency matrix of a directed n_1-cycle, which completes the proof. □

4.2. Kronecker product of stochastic matrices

Given stochastic matrices A and B, $A \otimes B$ is also stochastic where $\otimes$ denotes the Kronecker product. We will provide an expression of Kemeny's constant of $A \otimes B$ after applying Theorem 3.3.

Partition $P = A \otimes B$ as follows:

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \left[\begin{array}{c|ccc} a_{11}B & a_{12}B & \ldots & a_{1n}B \\ \hline a_{21}B & & & \\ \vdots & & A_1 \otimes B & \\ a_{n1}B & & & \end{array}\right]. \tag{33}$$

We use $\mathbf{e}_i$ to denote the vector with a 1 in the ith entry and zeros elsewhere. We have the following

Proposition 4.7. *Let $A = [a_{ij}]$ and B be stochastic matrices with stationary distribution vectors $\mathbf{x} = (x_i)$ and $\mathbf{y}$, respectively. Let $P = A \otimes B$. Then*

$$\kappa(P) = \kappa(P_1) + \kappa(P_2) + \frac{1}{1 - x_1}(\mathbf{e}_1^T (I - A + \mathbf{1}\,\mathbf{x}^T)^{-1}\mathbf{e}_1 - x_1),$$

where P_1 and P_2 are the stochastic complements in (33) of P_{11} and P_{22}, respectively.

Proof. Let $A = [a_{ij}]$ and B be matrices of size $n \times n$ and $m \times m$, respectively. It suffices to provide an explicit expression for γ in Theorem 3.3.

Let $\mathbf{x} = (x_i) \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$ be the stationary distribution vector for A and B, respectively. Then $\mathbf{x}^T A = \mathbf{x}^T$, $\mathbf{y}^T B = \mathbf{y}^T$ and $\mathbf{x}^T \mathbf{1} = \mathbf{y}^T \mathbf{1} = 1$. Let $\boldsymbol{\pi}^T = \begin{bmatrix} \boldsymbol{\pi}_1^T & \boldsymbol{\pi}_2^T \end{bmatrix}$ be conformally partitioned with P and be the stationary distribution vector for P. Since $\boldsymbol{\pi} = \mathbf{x} \otimes \mathbf{y}$, we have $\boldsymbol{\pi}_1 = x_1 \mathbf{y}$ and $\boldsymbol{\pi}_2 = (x_2, \dots, x_n) \otimes \mathbf{y}$. So, $\|\boldsymbol{\pi}_1\| = x_1$ and $\|\boldsymbol{\pi}_2\| = 1 - x_1$.

Set $q = mn - m$. Consider

$$\begin{bmatrix} \|\boldsymbol{\pi}_2\| \, \mathbf{1}_m \\ -\|\boldsymbol{\pi}_1\| \, \mathbf{1}_q \end{bmatrix} = -\|\boldsymbol{\pi}_1\| \begin{bmatrix} \mathbf{1}_m \\ \mathbf{1}_q \end{bmatrix} + \begin{bmatrix} \mathbf{1}_m \\ \mathbf{0}_q \end{bmatrix}.$$

Then

$$(I - P + \mathbf{1}\,\boldsymbol{\pi}^T)^{-1} \begin{bmatrix} \|\boldsymbol{\pi}_2\| \, \mathbf{1}_m \\ -\|\boldsymbol{\pi}_1\| \, \mathbf{1}_q \end{bmatrix} = -\|\boldsymbol{\pi}_1\| \begin{bmatrix} \mathbf{1}_m \\ \mathbf{1}_q \end{bmatrix} + (I - P + \mathbf{1}\,\boldsymbol{\pi}^T)^{-1}(\mathbf{e}_1 \otimes \mathbf{1}_m). \quad (34)$$

We claim that

$$(I - P + \mathbf{1}\,\boldsymbol{\pi}^T)^{-1}(\mathbf{e}_1 \otimes \mathbf{1}_m) = \mathbf{z} \otimes \mathbf{1}_m$$

where $\mathbf{z} = (I - A + \mathbf{1}_n \, \mathbf{x}^T)^{-1}\mathbf{e}_1$. Consider

$$(I - A \otimes B + \mathbf{1}(\mathbf{x} \otimes \mathbf{y})^T)(\mathbf{z} \otimes \mathbf{1}_m) = \mathbf{e}_1 \otimes \mathbf{1}_m \, .$$

This system has the form

$$(\mathbf{z} \otimes \mathbf{1}_m) - (A \otimes B)(\mathbf{z} \otimes \mathbf{1}_m) + (\mathbf{1}_n \otimes \mathbf{1}_m)(\mathbf{x} \otimes \mathbf{y})^T(\mathbf{z} \otimes \mathbf{1}_m) = (\mathbf{e}_1 \otimes \mathbf{1}_m).$$

This implies that

$$(\mathbf{z} - A\mathbf{z} + (\mathbf{1}_n \, \mathbf{x}^T)\mathbf{z} - \mathbf{e}_1) \otimes \mathbf{1}_m = \mathbf{0}.$$

Hence, our desired claim is established.

Now, applying (34) together with the claim to the expression of γ given in Theorem 3.3, we obtain

$$\begin{aligned}
\gamma &= \begin{bmatrix} \widehat{\boldsymbol{\pi}}_1^T & -\widehat{\boldsymbol{\pi}}_2^T \end{bmatrix} (I - P + \mathbf{1}\,\boldsymbol{\pi}^T)^{-1} \begin{bmatrix} \|\boldsymbol{\pi}_2\| \, \mathbf{1}_m \\ -\|\boldsymbol{\pi}_1\| \, \mathbf{1}_q \end{bmatrix} \\
&= \begin{bmatrix} \widehat{\boldsymbol{\pi}}_1^T & -\widehat{\boldsymbol{\pi}}_2^T \end{bmatrix} (-\|\boldsymbol{\pi}_1\| \, \mathbf{1} + \mathbf{z} \otimes \mathbf{1}_m) \\
&= (\begin{bmatrix} 1 & -\frac{1}{1 - \|\boldsymbol{\pi}_1\|}(x_2, \dots, x_n) \end{bmatrix} \otimes \mathbf{y}^T)(\mathbf{z} \otimes \mathbf{1}_m) \\
&= \frac{1}{1 - x_1} \, [1 - x_1 \quad -x_2 \quad \dots \quad -x_n] \, \mathbf{z} \\
&= \frac{1}{1 - x_1}(\mathbf{e}_1 - \mathbf{x})^T (I - A + \mathbf{1}_n \, \mathbf{x}^T)^{-1}\mathbf{e}_1 \\
&= \frac{1}{1 - x_1}(\mathbf{e}_1^T (I - A + \mathbf{1}_n \, \mathbf{x}^T)^{-1}\mathbf{e}_1 - x_1).
\end{aligned}$$

Therefore, we obtain the desired result. □

4.3. Sub-stochastic matrices with constant row sums

Here is the result of this subsection.

Proposition 4.8. *Let*

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix},$$

be an irreducible stochastic matrix, where $P_{11}\,\mathbf{1} = r_1\,\mathbf{1}$ *and* $P_{22}\,\mathbf{1} = r_2\,\mathbf{1}$ *for some* $0 \le r_1, r_2 < 1$. *Denote by* P_1 *and* P_2 *the stochastic complements of* P_{11} *and* P_{22}, *respectively. Then*

$$\|\boldsymbol{\pi}_1\| = \frac{1-r_2}{2-r_1-r_2}, \qquad \|\boldsymbol{\pi}_2\| = \frac{1-r_1}{2-r_1-r_2},$$
$$\kappa(P) = \kappa(P_1) + \kappa(P_2) + \frac{1}{2-r_1-r_2}.$$

Proof. Note that $P_{12}\,\mathbf{1} = (1-r_1)\,\mathbf{1}$ and $P_{21}\,\mathbf{1} = (1-r_2)\,\mathbf{1}$. From (19),

$$\begin{aligned} \theta &= \hat{\boldsymbol{\pi}}_2^T (I-P_{22})^{-1}\,\mathbf{1} + \hat{\boldsymbol{\pi}}_2^T (I-P_{22})^{-1} P_{21} (I - P_1 + \mathbf{1}\,\hat{\boldsymbol{\pi}}_1^T)^{-1} P_{12} (I-P_{22})^{-1}\,\mathbf{1} \\ &= \frac{1}{1-r_2} + \frac{1-r_1}{1-r_2} = \frac{2-r_1}{1-r_2}. \end{aligned}$$

On the other hand, we find from (4) that the aggregated matrix S is given by

$$S = \begin{bmatrix} r_1 & 1-r_1 \\ 1-r_2 & r_2 \end{bmatrix},$$

and so $\alpha_1 = \|\boldsymbol{\pi}_1\| = (1-r_2)/(2-r_1-r_2)$, $\alpha_2 = \|\boldsymbol{\pi}_2\| = (1-r_1)/(2-r_1-r_2)$, and $\|\boldsymbol{\pi}_1\|(1+\theta) = (3-r_1-r_2)/(2-r_1-r_2)$. Whence, in view of (15) and (18), we deduce that

$$\kappa(P) = \kappa(P_1) + \kappa(P_2) + \frac{1}{2-r_1-r_2}. \quad \square$$

5. Some bounds

Assume we are given the values of $\kappa(P_1)$ and $\kappa(P_2)$. According to Equation (15), we can provide lower and upper bounds to the value of $\kappa(P)$ once we are given bounds to the constant γ. In order to do this, we need to determine upper and lower bounds to the value of $\|\boldsymbol{\pi}_1\|$.

From the equation $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$ we find that

$$\boldsymbol{\pi}_1^T = \boldsymbol{\pi}_2^T P_{21}(I - P_{11})^{-1}, \qquad \boldsymbol{\pi}_2^T = \boldsymbol{\pi}_1^T P_{12}(I - P_{22})^{-1}.$$

Taking the infinity norms of both sides and using the identity $\|\mathbf{x}^T\|_\infty = \|\mathbf{x}\|_1$, we obtain

$$\|\boldsymbol{\pi}_1\| \le \|\boldsymbol{\pi}_2\| \, \|P_{21}\|_\infty \|(I - P_{11})^{-1}\|_\infty \le \|\boldsymbol{\pi}_2\| \frac{\|P_{21}\|_\infty}{1 - \|P_{11}\|_\infty},$$

where the latter inequality is valid if $\|P_{11}\|_\infty < 1$. A similar inequality can be obtained for $\|\boldsymbol{\pi}_2\|$. Combining both inequalities, under the assumption $\|P_{11}\|_\infty, \|P_{22}\|_\infty < 1$, we get

$$\frac{1 - \|P_{22}\|_\infty}{\|P_{12}\|_\infty} \le \frac{\|\boldsymbol{\pi}_1\|}{\|\boldsymbol{\pi}_2\|} \le \frac{\|P_{21}\|_\infty}{1 - \|P_{11}\|_\infty}.$$

Moreover, since $\|\boldsymbol{\pi}_2\| = 1 - \|\boldsymbol{\pi}_1\|$, we obtain

$$\frac{1 - \|P_{22}\|_\infty}{1 - \|P_{22}\|_\infty + \|P_{12}\|_\infty} \le \|\boldsymbol{\pi}_1\| \le \frac{\|P_{21}\|_\infty}{1 - \|P_{11}\|_\infty + \|P_{21}\|_\infty}. \tag{35}$$

Since the matrix P is known, the above bounds can be actually computed at a low computational cost.

Now a bound to the constant γ can be obtained by relying on Equation (18) coupled with (19) (one can use (21) or (20)). From (18), $\gamma = (1 + \theta)\|\boldsymbol{\pi}_1\| - 1$. We see from (35) that, if $1 + \theta \ge 0$, then

$$(\theta + 1)\frac{1 - \|P_{22}\|_\infty}{1 - \|P_{22}\|_\infty + \|P_{12}\|_\infty} - 1 \le \gamma \le (\theta + 1)\frac{\|P_{21}\|_\infty}{1 - \|P_{11}\|_\infty + \|P_{21}\|_\infty} - 1.$$

A similar inequality holds if $1 + \theta < 0$.

Note that $\|(I - P_{22})^{-1}P_{21}\|_\infty = 1$. Concerning θ, it follows from (19) that

$$\theta \le \; \|(I - P_{22})^{-1}\|_\infty \left(1 + \|P_{12}\|_\infty \|(I - P_1 + \mathbf{1}\,\widehat{\boldsymbol{\pi}}_1^T)^{-1}\|_\infty\right). \tag{36}$$

The upper bound to γ is expressed in terms of $(I - P_{22})^{-1}$, $(I - P_1 + \mathbf{1}\,\widehat{\boldsymbol{\pi}}_1^T)^{-1}$ and the norms of the blocks P_{ij}.

For another bound on γ, consider the expression of γ in Theorem 3.3, *i.e.*,

$$\gamma = \begin{bmatrix} \widehat{\boldsymbol{\pi}}_1^T & -\widehat{\boldsymbol{\pi}}_2^T \end{bmatrix} (I - P + \mathbf{1}\,\boldsymbol{\pi}^T)^{-1} \begin{bmatrix} \|\boldsymbol{\pi}_2\|\,\mathbf{1} \\ -\|\boldsymbol{\pi}_1\|\,\mathbf{1} \end{bmatrix}.$$

Taking the infinity norm of both sides in the above equation yields

$$|\gamma| \le 2\|(I - P + \mathbf{1}\,\boldsymbol{\pi}^T)^{-1}\|_\infty \max(\|\boldsymbol{\pi}_1\|, 1 - \|\boldsymbol{\pi}_1\|).$$

5.1. A perturbation result

Let P be an $n \times n$ stochastic and irreducible matrix, and E be $n \times n$ matrix such that $E\,\mathbf{1} = 0$. Let $\epsilon > 0$. Then $(P + \epsilon E)\,\mathbf{1} = \mathbf{1}$. Assume that $P(\epsilon) := P + \epsilon E$ is stochastic and $\|E\|_\infty \le 1$.

Here, our goal is to relate $\kappa(P(\epsilon))$ and $\kappa(P)$. We have

$$\kappa(P(\epsilon)) = \operatorname{Tr}((I - P(\epsilon) + \mathbf{1}\,\mathbf{h}^T)^{-1}), \qquad \kappa(P) = \operatorname{Tr}((I - P + \mathbf{1}\,\mathbf{h}^T)^{-1}),$$

where $\mathbf{h}$ is any vector such that $\mathbf{h}^T\,\mathbf{1} = 1$. By subtracting both sides of the above equations we get

$$\kappa(P(\epsilon)) - \kappa(P) = \operatorname{Tr}((I - P(\epsilon) + \mathbf{1}\,\mathbf{h}^T)^{-1}(P(\epsilon) - P)(I - P + \mathbf{1}\,\mathbf{h}^T)^{-1}).$$

Neglecting $O(\epsilon^2)$ terms, we obtain

$$\begin{aligned} \kappa(P(\epsilon)) - \kappa(P) &= \epsilon \operatorname{Tr}((I - P + \mathbf{1}\,\mathbf{h}^T)^{-1} E (I - P + \mathbf{1}\,\mathbf{h}^T)^{-1}) \\ &= \epsilon \operatorname{Tr}((I - P + \mathbf{1}\,\mathbf{h}^T)^{-2} E) + O(\epsilon^2). \end{aligned}$$

This estimate can be also deduced as a specific case of [7, Lemma 3.2].

From the Cauchy–Schwarz inequality, we have $\operatorname{Tr}(AB) \le \|A\|_F \|B\|_F$, where $\|\cdot\|_F$ is the Frobenius norm. It follows that

$$|\kappa(P(\epsilon)) - \kappa(P)| \le \epsilon \|(I - P + \mathbf{1}\,\mathbf{h}^T)^{-2}\|_F \|E\|_F + O(\epsilon^2). \tag{37}$$

Combining the above bound with Example 4.4 yields the following result that concerns Kemeny's constant of a matrix associated with an almost bipartite graph.

Corollary 5.1. *Let P be the stochastic matrix defined in Example 4.4 and E be a matrix such that $E\,\mathbf{1} = 0$ and $\|E\|_\infty \le 1$. Set $P(\epsilon) = P + \epsilon E$ and assume that $P(\epsilon)$ is stochastic in a neighborhood of 0. Then, up to within $O(\epsilon^2)$ terms we have*

$$\left| \kappa(P(\epsilon)) - \kappa(P_1) - \kappa(P_2) - \frac{1}{2} \right| \le \epsilon \|(I - P + \mathbf{1}\,\mathbf{h}^T)^{-2}\|_F \|E\|_F + O(\epsilon^2),$$

where $\mathbf{h}$ is any vector such that $\mathbf{h}^T\,\mathbf{1} = 1$

Similar bounds can be obtained for the "stochastic" perturbation of a cyclic matrix having arbitrary cyclicity index and for the perturbation of the Kronecker product of two stochastic matrices.

Analogously, we may provide a "perturbed version" of Proposition 4.8 obtained by combining Proposition 4.8 itself with Equation (37).

6. A divide-and-conquer algorithm

We can use the analysis done in Section 3 to construct a divide-and-conquer algorithm for computing Kemeny's constant of a stochastic sparse matrix P. In fact, it is sufficient to identify a partitioning of the form (1) to start a recursion procedure. This procedure employs Theorem 3.3 to express $\kappa(P)$ in terms of $\kappa(P_1)$ and $\kappa(P_2)$ of the censored chains and the γ, where P_1 and P_2 are the stochastic complements. Algorithm 1 compactly describes the entire procedure in a recursive formulation.

Algorithm 1: Divide-and-conquer algorithm for the computation of Kemeny's constant.

1 Kemeny(P,$\boldsymbol{\pi}$) /* Implementation as a recursive function */
  Input: Stochastic matrix P, stationary distribution vector $\boldsymbol{\pi}$, an integer $n_0 > 0$.
  Output: Kemeny's constant $\kappa(P)$ of P.
2 /* In the following the $\leftarrow$ means variable assignment while the $=$ are used to define shorthand. */
3 $n \leftarrow \text{size}(P)$;
4 **if** $n < n_0$ **then**
5   $\kappa(P) \leftarrow \text{Tr}\left((I - P + \mathbf{1}\mathbf{1}^T/n)^{-1}\right) - 1$;
6   **return**
7 **else**
8   $m \leftarrow \lfloor n/2 \rfloor$;
9   $P_{11} = P(1:m, 1:m)$; /* Block structure */
10   $P_{12} = P(1:m, m+1:n)$;
11   $P_{21} = P(m+1:n, 1:m)$;
12   $P_{22} = P(m+1:n, m+1:n)$;
13   $\boldsymbol{\pi}_1 = \boldsymbol{\pi}(1:m)$;
14   $\boldsymbol{\pi}_2 = \boldsymbol{\pi}(m+1:n)$;
15   $\hat{\boldsymbol{\pi}}_i \leftarrow \frac{1}{\|\boldsymbol{\pi}_i\|}\boldsymbol{\pi}_i, \quad i = 1, 2$;
16   /* Remark: we can compute the LU factorization of $I - P_{22}$ once and reuse it. */
17   $P_1 \leftarrow P_{11} + P_{12}(I - P_{22})^{-1}P_{21}$; /* stochastic complements */
18   $P_2 \leftarrow P_{22} + P_{21}(I - P_{11})^{-1}P_{12}$;
19   $\mathbf{x} \leftarrow (I - P_{22})^{-1}\mathbf{1}$; /* Computation of θ using (19) */
20   $\mathbf{y} \leftarrow (I - P_1 + \mathbf{1}\,\hat{\boldsymbol{\pi}}_1^T)^{-1}(P_{12}\mathbf{x})$;
21   $\mathbf{y} \leftarrow (I - P_{22})^{-1}(P_{21}\mathbf{y})$;
22   $\theta \leftarrow \hat{\boldsymbol{\pi}}_2^T(\mathbf{x} + \mathbf{y})$;
23   $\gamma \leftarrow \|\boldsymbol{\pi}_1\|\theta - \|\boldsymbol{\pi}_2\|$; /* Compute correction as in (18) */
24   $\kappa(P_1) \leftarrow$ Kemeny(P_1,$\hat{\boldsymbol{\pi}}_1$); /* Recursion */
25   $\kappa(P_2) \leftarrow$ Kemeny(P_2,$\hat{\boldsymbol{\pi}}_2$); /* Recursion */
26   $\kappa(P) \leftarrow \kappa(P_1) + \kappa(P_2) + \gamma$;
27   **return**

To effectively apply this strategy, it is crucial to efficiently address several computational subproblems. Let us set aside the computation of the stationary distribution vector $\boldsymbol{\pi}$ for the starting chain P, which is the essential component required to initiate the entire procedure (to this regard, for large-scale problems, the algorithms proposed in [13] might be used). The most significant part of the computation lies in the solution of the linear systems to lines 17–21 of Algorithm 1.

Let us focus on the solution of the following systems:

$$(I - P_{ii})\mathbf{x} = \mathbf{b}, \qquad i = 1, 2,$$

where we further assume that the block matrices $\{P_{ii}\}_{i=1}^{2}$ can benefit from sparse storage. By construction, the blocks P_{ii}, $i = 1, 2$ are sub-stochastic, i.e., $(P_{ii})_{p,q} \geq 0$ for all p, q, and $P_{ii}\,\mathbf{1} \leq \mathbf{1}$, $P_{ii}\,\mathbf{1} \neq \mathbf{1}$. Since P is irreducible, the matrices $I - P_{ii}$, $i = 1, 2$, are non-singular M-matrices. This property allows us to resort to different efficient iterative strategies for the solution of the systems involved. Since in general matrices will be non-symmetric, one can consider using GMRES [33, Section 6.5] or BiCGstab [33, Section 7.4.2] for solving the different linear systems. (In any case it will be necessary to have a preconditioner available to accelerate the convergence of the Krylov method in question.) Since we are working with M-matrices, a natural choice is to use incomplete factorizations. Specifically, we can use Incomplete LU factorizations (ILU), that is, we can approximate the matrices as

$$I - P_{ii} = \widetilde{L}_i \widetilde{U}_i + R_i \approx \widetilde{L}_i \widetilde{U}_i \qquad i = 1, 2,$$

with the residual matrix R_i satisfying certain constraints, such as having zero entries in some prescribed locations–either static, determined on the base of the natural occurring fill-in during the computation, or via thresholding on their entries. For M-matrices, the existence of such objects is guaranteed by the fact that Gaussian elimination and non-diagonal dropping of the entries preserves the property of being a non singular M-matrix–see [15] for the original proof or [33, Theorem 10.1] for a modern explanation. Similarly, to precondition the system on the line 20 of the algorithm that contains the matrix P_1, we can consider an incomplete factorization of an approximation of the matrix $I - P_1$, namely,

$$\begin{aligned} I - P_1 &= I - P_{11} - P_{12}(I - P_{22})^{-1} P_{21} \\ &\approx I - P_{11} - P_{12}(\operatorname{diag}(I - P_{22}))^{-1} P_{21} \approx \widetilde{L}_3 \widetilde{U}_3, \end{aligned}$$

where $\operatorname{diag}(I - P_{22})$ is the diagonal matrix formed by the diagonal entries of $I - P_{22}$. It is known that the computation of LU factorization, incomplete or not, benefits from the reordering of the entries of the matrix itself, see, e.g., [2]. Furthermore, as seen in Proposition 4.8, the closer a matrix is to block diagonal after appropriate permutation, the easier the global calculation will be. Since the off-diagonal matrices in the block decomposition (1) will be made up of a few nonzero elements, the natural choice for the permutation algorithm is to use Nested Dissection permutation algorithm [20]; an example of the application of this algorithm is given in Fig. 1, from which we observe that the resulting matrix has the desired “quasi block-diagonal” structure. We want to underline that this initial permutation step would also be advisable if one wants to obtain Kemeny’s constant directly from the expression (6), since the related effect of this choice is also that of a permutation which is fill-reducing for the computation of LU factorization of sparse matrices, a step needed for the efficient computation of the matrix inverse in (6).

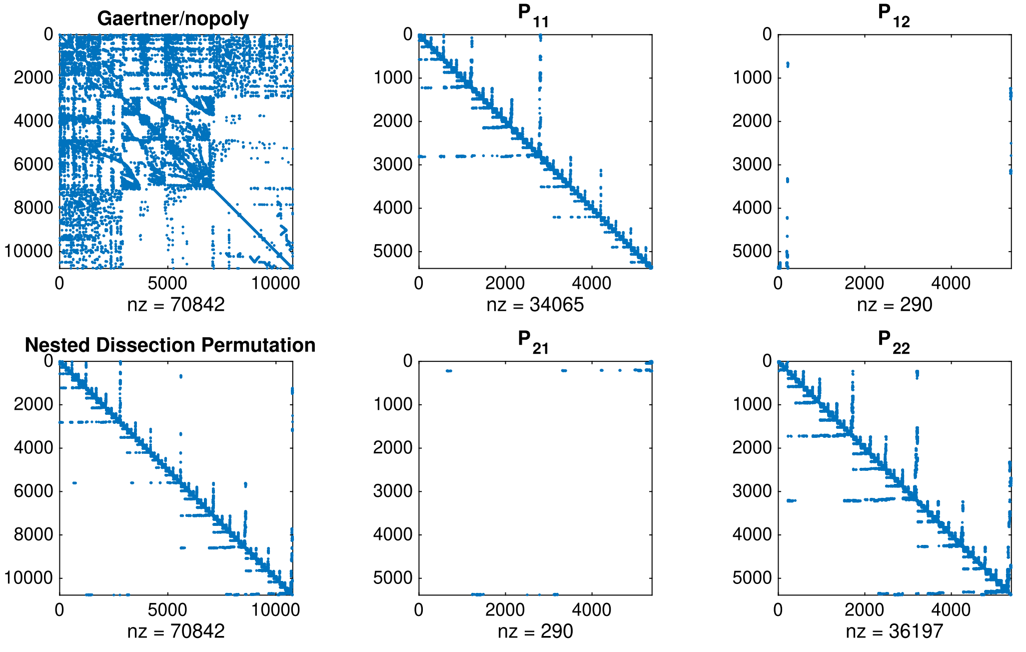

Fig. 1. Application of the Nested Dissection algorithm from [20] to the `Gaertner/nopoly` matrix from the `SuiteSparse Matrix Collection` (formerly the University of Florida Sparse Matrix Collection) [12].

6.1. Low-precision randomized approximation

We note that since we want to solve intermediate linear systems using an iterative method, what we are actually calculating is an approximation of Kemeny's constant. For this reason, it makes sense to consider randomized algorithms for the direct approximation of (6) through the trace of a matrix. Such approaches have been also used for instance in [30], for reversible Markov chains, and in [35], for Markov chains modeling a random walk on an undirected graph. We consider here the application of the Hutch++ algorithm [32], which is a straightforward improvement on the Hutchinson estimator [18] used in [35]. In our case this means having an *oracle* that computes

$$\mathbf{y} = (I - P + \mathbf{1}\,\mathbf{h}^T)^{-1}\mathbf{x} \equiv A\mathbf{x}, \quad \mathbf{h} = \mathbf{1}/n. \tag{38}$$

Then the following results give us the number of oracle calls, i.e., linear system solutions, we have to approximate $\operatorname{Tr}\left((I - P + \mathbf{1}\,\mathbf{h}^T)^{-1}\right)$ within a given tolerance.

Theorem 6.1 *([32, Theorem 1]). If Hutch++ is implemented with* $\ell = O(\sqrt{\log(1/\delta)}/\epsilon + \log(1/\delta))$ *matrix-vector multiplication queries, then for any positive semidefinite matrix* A*, with probability* $\geq 1 - \delta$*, the output of Hutch++(*A*) satisfies*

$$(1 - \varepsilon)\operatorname{Tr}(A) \leq \mathit{Hutch{+}{+}}(A) \leq (1 + \varepsilon)\operatorname{Tr}(A).$$

Similarly to what was done for the solution of linear systems in the divide-and-conquer algorithm, we can use the PCG with an Incomplete Cholesky preconditioner calculated on $I - P$, i.e.,

$$A \approx \widetilde{L}^{-T} \left(I - \frac{1}{1 + (\mathbf{h}^T \widetilde{L}^{-T})(\widetilde{L}^{-1}\mathbf{1})} (\widetilde{L}^{-1}\mathbf{1})(\mathbf{h}^T \widetilde{L}^{-T}) \right) \widetilde{L}^{-1},$$

as an oracle for the calculation of the products $A\mathbf{x}$ necessary for Hutch++. We point out that the Hutch++ algorithm works under the assumption that the matrix of which we approximate the trace is symmetric positive semidefinite, and this assumption is verified for Markov chains modeling a random walk on an undirected graph.

6.2. Numerical examples

This section contains some numerical examples in which the performance of the algorithms obtained starting from the theoretical analysis is analyzed. All experiments are reproducible starting from the code contained in the repository github.com/Cirdans-Home/Kemeny-and-Conquer. All experiments are performed on a vertex of the Toeplitz cluster at the University of Pisa equipped with an Intel® Xeon® CPU E5-2650 v4 at 2.20 GHz and 250 Gb of RAM, using MATLAB 9.10.0.1602886 (R2021a). The Markov chains used in the examples are built employing matrices from the `SuiteSparse Matrix Collection` (formerly the University of Florida Sparse Matrix Collection) [12]. Specifically, we build the probability transition matrix P from irreducible adjacency matrices A as

$$P = \mathrm{diag}(\hat{A}\,\mathbf{1})^{-1}\hat{A},$$

for $\hat{A}$ the matrix obtained from A with all weights set to 1. For the low-precision randomized case we use instead

$$P = \mathrm{diag}(\hat{A}\,\mathbf{1})^{-1/2}\hat{A}\,\mathrm{diag}(\hat{A}\,\mathbf{1})^{-1/2},$$

for cases with $A = A^T$, i.e., the graph is undirected. All the relative errors in the following experiments are computed with respect to Kemeny's constant obtained directly, i.e., applying (6) by computing the whole matrix inverse.

6.2.1. Low-precision randomized approximation

We first focus on using randomized estimators from Section 6.1 for the trace of a matrix. Table 1 contains the estimates obtained using Hutch++ with the parameters (δ, ϵ) of Theorem 6.1 chosen as $\delta = 1/4$, and $\epsilon = 10^{-1}$, i.e., we use a set of $l = 13$ random sample vectors. As an internal solver for the oracle calculation we use the PCG preconditioned with an ICHOL(0), i.e., such as to preserve the sparsity pattern of the

Table 1
Randomized approximation of Kemeny's constant with Hutch++algorithm. The relative error and the timings (s) for the Hutch++is the average over 100 repetitions. The inner iterative solver is PCG preconditioned by the ICHOL(0)-based preconditioner from 6.1 and with a tolerance of 10^{-2}. The number of samples is $l = 13$, that corresponds to $\delta = 1/4$ and $\epsilon = 10^{-1}$.

Matrix	n	Time (s)		Rel. Error
		Direct	Hutch++	
`Pajek/USpowerGrid`	4941	2.11	0.43	2.69e-04
`Gaertner/nopoly`	10774	21.97	2.64	7.58e-03
`Gleich/minnesota`	2640	0.33	0.15	1.94e-03

starting matrix in $\widetilde{L}$. Since the external precision we expect to achieve is of the order of 10^{-2}, the tolerance for the iterative method is chosen to be 10^{-3}. In all cases, we always start from the matrix reordered with the Nested Dissection permutation algorithm [20].

6.2.2. Divide-and-conquer algorithm

Here we test the application of the divide-and-conquer strategy on the transition matrices built from the matrices in `SuiteSparse Matrix Collection` [12] in the previous section. We use two variants of Algorithm 1: one called Recursive in Table 2 exploits an incomplete ILU(0) factorization and the GMRES to solve systems with left term $I - P_{22}$, and the other called direct-recursive exploits the LU factorization of the matrix, already needed for the calculation of the stochastic complement, also for the solution of the two auxiliary systems with the same matrix for the calculation of θ. From the results in Table 2 we observe that in most cases the divide-and-conquer algorithm manages to reduce the computational time compared to the direct computation of Kemeny's constant. In some cases we observe the absence of an improvement, investigating in detail what we observe is that the decomposition into blocks done by halving is far from being optimal. This causes both the creation of denser stochastic complements and higher solution times for auxiliary linear systems. The implementation of nested dissection in MATLAB does not have in output the limitation of the *clusters* obtained, being able to use that should increase the advantage of the recursive version compared to the one in which the recursion is done by simple halving.

7. Conclusions

Kemeny's constant $\kappa(P)$ of a stochastic matrix P has been expressed in terms of Kemeny's constants of the stochastic complements obtained from a block partitioning, and the constant γ. Explicit expressions of $\kappa(P)$ have been provided for the transition matrix of a periodic Markov chain, the Kronecker product of stochastic matrices, and sub-stochastic matrices with constant row sums. The main result, Theorem 3.3, has been used to design a divide-and-conquer algorithm for recursively computing $\kappa(P)$. Numerical experiments applied to real-world graphs show the effectiveness of this approach especially in the case of nearly completely decomposable matrices.

Table 2

Performance of the recursive implementation of the divide-and-conquer algorithm for computing Kemeny's constant on some test matrices. If the matrix name has a † symbol, then the experiment has been run on the largest connected component of the graph, i.e., on the largest irreducible sub-chain.

Matrix	n	$\kappa(P)$	Time (s)			Rel. Error	
			Direct	Recursive	Dir-Rec	Recursive	Dir-Rec
`Gaertner/big`	13209	58134.53	17	5.81	**4.15**	1.57e-08	4.09e-09
`vanHeukelum/cage10`	11397	15378.12	**11.06**	44.83	56.76	8.28e-13	1.25e-11
`vanHeukelum/cage11`	39082	51177.08	**315.27**	1259.11	1591.04	2.59e-10	1.53e-10
`HB/gre_1107`	1107	1483.57	**0.04**	0.10	0.11	1.97e-09	1.09e-09
`Gaertner/nopoly`	10774	171656.87	9.56	9.14	**6.71**	3.25e-07	3.09e-07
`Gaertner/pesa`	11738	131250.78	11.45	6.36	**3.30**	2.15e-07	2.28e-07
`Gleich/usroads-48`	126146	1818057.53	8243.57	743.37	**542.88**	8.65e-07	1.60e-07
`Barabasi/NotreDame_www`†	34643	1173610.94	172.36	94.28	**93.20**	2.85e-04	3.58e-05
`Pajek/USpowerGrid`	4941	30166.55	1.32	1.48	**0.58**	3.01e-08	2.80e-08
`Gleich/minnesota`†	2640	18243.53	0.30	0.34	**0.22**	8.37e-08	3.53e-08

As Kemeny's constant measures the expected time of a Markov chain X to travel between two randomly chosen states, a natural question arises: "What interpretation can be ascribed to Kemeny's constant of a censored Markov chain X_1 associated with the original Markov chain X?". Since X_1 is induced from X, it would be interesting to provide insights into what features of X are captured in Kemeny's constant of X_1. Partitioning the state space of X into two subsets yields two censored Markov chains. We can see from Theorem 3.3 that Kemeny's constants of these censored Markov chains are interdependent. Understanding how partitioning the state space influences Kemeny's constants of censored Markov chains would be beneficial for gaining insights into the question.

Declaration of competing interest

There is no competing interest.

Data availability

Online git repository.

Acknowledgements

The authors sincerely thank the anonymous referee for the thorough review and valuable suggestions, which have significantly enhanced the clarity and quality of the presentation.

References

[1] D. Altafini, D.A. Bini, V. Cutini, B. Meini, F. Poloni, An edge centrality measure based on the Kemeny constant, SIAM J. Matrix Anal. Appl. 44 (2) (2023) 648–669.
[2] M. Benzi, D.B. Szyld, A. van Duin, Orderings for incomplete factorization preconditioning of nonsymmetric problems, SIAM J. Sci. Comput. 20 (5) (1999) 1652–1670.
[3] D. Bini, J.J. Hunter, G. Latouche, B. Meini, P. Taylor, Why is Kemeny's constant a constant?, J. Appl. Probab. 55 (4) (2018) 1025–1036.
[4] A. Brauer, Limits for the characteristic roots of a matrix. VII, Duke Math. J. 25 (1958) 583–590.
[5] J. Breen, S. Butler, N. Day, C. DeArmond, K. Lorenzen, H. Qian, J. Riesen, Computing Kemeny's constant for barbell-type graphs, Electron. J. Linear Algebra 35 (2019) 583–598.
[6] J. Breen, N. Faught, C. Glover, M. Kempton, A. Knudson, A. Oveson, Kemeny's constant for nonbacktracking random walks, Random Struct. Algorithms 63 (2) (2023) 343–363.
[7] J. Breen, S. Kirkland, A structured condition number for Kemeny's constant, SIAM J. Matrix Anal. Appl. 40 (4) (2019) 1555–1578.
[8] P. Brémaud, Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues, vol. 31, Springer Science & Business Media, 2001.
[9] L. Ciardo, The Braess' paradox for pendent twins, Linear Algebra Appl. 590 (2020) 304–316.
[10] L. Ciardo, G. Dahl, S. Kirkland, On Kemeny's constant for trees with fixed order and diameter, Linear Multilinear Algebra 70 (12) (2022) 2331–2353.
[11] E. Crisostomi, S. Kirkland, R. Shorten, A Google-like model of road network dynamics and its application to regulation and control, Int. J. Control 84 (3) (2011) 633–651.

[12] T.A. Davis, Y. Hu, The University of Florida sparse matrix collection, ACM Trans. Math. Softw. 38 (1) (2011) 1.
[13] H. De Sterck, K. Miller, E. Treister, I. Yavneh, Fast multilevel methods for Markov chains, Numer. Linear Algebra Appl. 18 (6) (2011) 961–980.
[14] P.G. Doyle, The Kemeny constant of a Markov chain, arXiv preprint, arXiv:0909.2636, 2009.
[15] K. Fan, Note on M-matrices, Quart. J. Math. Oxford Ser. (2) 11 (1960) 43–49.
[16] N. Faught, M. Kempton, A. Knudson, A 1-separation formula for the graph Kemeny constant and Braess edges, J. Math. Chem. 60 (1) (2022) 49–69.
[17] R.A. Horn, C.R. Johnson, Matrix Analysis, second edition, Cambridge University Press, Cambridge, 2013.
[18] M.F. Hutchinson, A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines, Commun. Stat., Simul. Comput. 19 (2) (1990) 433–450.
[19] J. Jang, S. Kim, M. Song, Kemeny's constant and Wiener index on trees, Linear Algebra Appl. 674 (2023) 230–243.
[20] G. Karypis, V. Kumar, A fast and high quality multilevel scheme for partitioning irregular graphs, SIAM J. Sci. Comput. 20 (1) (1998) 359–392.
[21] J.G. Kemeny, Generalization of a fundamental matrix, Linear Algebra Appl. 38 (1981) 193–206.
[22] J.G. Kemeny, J.L. Snell, Finite Markov Chains, Undergraduate Texts in Mathematics, Springer-Verlag, New York-Heidelberg, 1976, Reprinting of the 1960 original.
[23] S. Kim, Families of graphs with twin pendent paths and the Braess edge, Electron. J. Linear Algebra (2022) 9–31.
[24] S. Kim, J. Breen, E. Dudkina, F. Poloni, E. Crisostomi, On the effectiveness of random walks for modeling epidemics on networks, PLoS ONE 18 (1) (2023) e0280277.
[25] S. Kim, N. Madras, A. Chan, M. Kempton, S. Kirkland, A. Knudson, Bounds on Kemeny's constant of a graph and the Nordhaus-Gaddum problem, arXiv preprint, arXiv:2309.05171, 2023.
[26] S. Kirkland, Fastest expected time to mixing for a Markov chain on a directed graph, Linear Algebra Appl. 433 (11–12) (2010) 1988–1996.
[27] S. Kirkland, Directed forests and the constancy of Kemeny's constant, J. Algebraic Comb. 53 (1) (2021) 81–84.
[28] S. Kirkland, Y. Li, J. McAlister, X. Zhang, Edge addition and the change in Kemeny's constant, arXiv preprint, arXiv:2306.04005, 2023.
[29] S. Kirkland, Z. Zeng, Kemeny's constant and an analogue of Braess' paradox for trees, Electron. J. Linear Algebra 31 (2016) 444–464.
[30] S. Li, X. Huang, C.-H. Lee, An efficient and scalable algorithm for estimating Kemeny's constant of a Markov chain on large graphs, in: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, KDD '21, Association for Computing Machinery, New York, NY, USA, 2021, pp. 964–974.
[31] C.D. Meyer, Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems, SIAM Rev. 31 (2) (1989) 240–272.
[32] R.A. Meyer, C. Musco, C. Musco, D.P. Woodruff, Hutch++: optimal stochastic trace estimation, in: Symposium on Simplicity in Algorithms (SOSA), Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2021, pp. 142–155.
[33] Y. Saad, Iterative Methods for Sparse Linear Systems, second edition, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2003.
[34] X. Wang, J.L. Dubbeldam, P. Van Mieghem, Kemeny's constant and the effective graph resistance, Linear Algebra Appl. 535 (2017) 231–244.
[35] W. Xu, Y. Sheng, Z. Zhang, H. Kan, Z. Zhang, Power-law graphs have minimal scaling of Kemeny constant for random walks, in: The Web Conference 2020 - Proceedings of the World Wide Web Conference, WWW 2020, 2020, pp. 46–56.
[36] S. Yilmaz, E. Dudkina, M. Bin, E. Crisostomi, P. Ferraro, R. Murray-Smith, T. Parisini, L. Stone, R. Shorten, Kemeny-based testing for COVID-19, PLoS ONE 15 (11) (2020) e0242401.