Coevolutionary Systems and PageRank

S.Y. Chong[a,b,*], P. Tiňo[b], J. He[c]

[a] School of Computer Science, University of Nottingham Malaysia Campus, 43500 Semenyih, Malaysia
[b] School of Computer Science, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
[c] School of Science and Technology, Nottingham Trent University, Nottingham NG11 8NS, UK

Abstract

Coevolutionary systems have been used successfully in various problem domains involving situations of strategic decision-making. Central to these systems is a mechanism whereby finite populations of agents compete for reproduction and adapt in response to their interaction outcomes. Outcomes from their behavioral interactions express preferences over the candidate solutions they implement in competitive settings. A recent framework for precise characterizations of competitive coevolutionary systems was introduced. Its two main features are: (1) A directed graph (digraph) representation that fully captures the underlying structure arising from pairwise preferences over solutions. (2) Coevolutionary processes are modelled as random walks on the digraph. Here, we study a deep connection between coevolutionary systems and PageRank, and develop a principled approach to measure and rank the performance (importance) of solutions (vertices) in coevolutionary digraphs. In PageRank formalism, B transfers part of its authority to A if A dominates B (there is an arc from B to A in the digraph), and so PageRank authority indicates the importance of a vertex. Upon suitable normalization, PageRank authorities have a natural interpretation of long-term visitation probabilities over the digraph by the coevolutionary random walk. We prove that PageRank for any coevolutionary system exists

*Corresponding author
Email addresses: Siang-Yew.Chong@nottingham.edu.my (S.Y. Chong), P.Tino@cs.bham.ac.uk (P. Tiňo), jun.he@ieee.org (J. He)

and derive closed-form expressions to calculate PageRank authorities. Changes to PageRank authorities due to changes in restart probability setting for any coevolutionary system are quantified precisely. Furthermore, PageRank authorities can be approximated effectively and our empirical studies show how they characterize coevolutionary digraphs with different underlying structures.
Keywords: Coevolutionary systems, PageRank, Markov chains

1. Introduction

Coevolutionary systems that are inspired by natural evolutionary processes have been applied extensively and shown remarkable success in problem areas involving situations of strategic decision-making. These include simulation tools involving a collection of interacting, adaptive agents to understand conditions for the emergence of complex, intelligent behaviors in the real-world [1, 2] and algorithms to generate high performance agents with minimal preprogrammed knowledge in competitive settings [3, 4].

Competitive coevolutionary systems share a common framework and are specified by: (1) the problem whereby the solutions are the set of all alternatives (strategies in a two-player game) with a pairwise preference relation indicating the superior alternative, and (2) the process whereby the finite population of agents (each implementing a strategy to play) goes through a process of selection and variation guided only by their interaction (game play) outcomes. This framework allows for various design choices and parameters in strategy representation (finite state machines, neural networks, etc.), variation approaches (recombination and mutation) to generate new distinct strategies, and selection operations that favours higher performing strategies for reproduction.

All coevolutionary systems implement a mechanism that requires the population of agents to compete for reproduction [5]. This can be double-edged in competitive settings. The system could exploit interactions to search increasingly superior strategies or succumbs to the deleterious effects from using relative fitness evaluations in the search process. These *pathologies* have been studied [6, 7, 8] and include formal methods [9, 10] characterizing cyclic coevolutionary dynamics due to the choice of selection mechanism in an evolutionary game theoretic setting in the framework of continuous dynamical systems.

Other studies have found that problem structures induced by pairwise relations can affect coevolutionary search [11]. A new framework that formally specifies and models coevolutionary systems is introduced [12]. It uses a digraph representation of coevolutionary problem where the vertex set corresponds to the strategy set and the arc set captures preferences over all pairs of strategies. Coevolutionary processes are modelled as discrete time Markov chains operating on the digraph. A distinct population-one coevolutionary algorithm corresponds to a specific implementation of random walk on digraphs. Other learning algorithms involving self play for any problem in the form of a two-player game can be modelled within this framework. In this manner, complete qualitative characterization of cycle structures underlying coevolutionary problems as well as quantitative characterizations of coevolutionary processes are obtained.

Performing quantitative characterizations on these coevolutionary Markov chains (CMCs) requires prior knowledge of the digraph's underlying structure. As in the case of evolutionary algorithms [13], the expected hitting times of the absorbing class for CMCs operating on reducible digraphs are studied in [12]. However, a CMC can operate on irreducible digraphs and so its stationary distribution is of interest. Can other useful quantitative characterizations of digraph structures commonly found in coevolutionary problems be obtained? We establish a direct connection between coevolutionary systems and large network analysis methodologies [14] particularly PageRank (used originally to measure webpage importance [15, 16]). In PageRank formalism, strategies transfer part of their authorities to those that dominate them at pairwise interactions level and these authorities give indication of strategies' performances. Furthermore, PageRank authorities with suitable normalization gives the long-term visitation probabilities of CMC with restart operating over the coevolutionary digraph.

This connection allows us to develop a principled approach to measure and rank the performance of individual strategies that correspond to the vertex set for any digraph representation of the coevolutionary problem. In Section 2, we establish several theoretical results supporting this approach. We prove that PageRank for any coevolutionary system exists, derive closed-form expressions to calculate PageRank authorities, and quantify their changes due to varying restart probability. Section 3 introduces a benchmark of coevolutionary digraphs having different structures that is used in the controlled empirical studies to demonstrate how PageRank authorities provide quantitative characterizations of the digraphs. Section 4 concludes with remarks for future studies.

2. Coevolutionary Systems and PageRank

The framework in [12] provides a formal approach to model and contruct coevolutionary systems as random walks on digraphs. We consider a wide range of problems modelled as two-player strategic games $C = (S, R)$ for which the full underlying structures are captured by coevolutionary digraphs $D_C = (V_S, A_R)$. V_S is a non-empty, finite vertex set that corresponds to the set of pure strategies S. $A_R \subset V_S \times V_S$ is a finite arc set of ordered pairs of distinct vertices that corresponds to the dominance relation R on S. The arc $(u, v) \in A_R$ (uv or $u \to v$) indicates v dominating u. Note that standard digraph terminology reverses our domination definition [17].

2.1. Structural Characterizations of Coevolutionary Digraphs

There are unique structures underlying coevolutionary digraphs D_C that allow us to characterize them. A D_C's underlying graph obtained by replacing arcs for all pairs of vertices by single edges, $UG(D_C) = (V_S, E)$, is complete. One characterization relates to edge biorientation, e.g., either the orientation uv or vu (win-lose) or both $\{uv, vu\} \in A_R$ (draw). Let $V_{S(n)}$ be the vertex set with $n = |V_{S(n)}|$ number of vertices. For $n \geq 2$, we obtain two coevolutionary digraph classes: (1) tournament $T(V_{S(n)}) \in \mathcal{T}(V_{S(n)})$ as orientations of $UG(D_C)$, and (2) semicomplete digraph $SD(V_{S(n)}) \in \mathcal{SD}(V_{S(n)})$ as biorientations of $UG(D_C)$ [17]. $\mathcal{T}(V_S)$ represents all coevolutionary games with win-lose outcomes and $\mathcal{SD}(V_S)$ generalizes further by allowing draws.

Further characterizations relate to global structures on reducibility (grouping of related vertices) and connectivity (reachability of vertices). A reducible $D_C(V_{S(n)})$ admits a vertex partition on $V_{S(n)}$ into two disjoint, nonempty subsets ($V_S^1 \cap V_S^2 = \emptyset$, $V_S^1 \cup V_S^2 = V_{S(n)}$) whereby $V_S^1 \mapsto V_S^2$ (each $v \in V_S^2$ dominates all $u \in V_S^1$ only). Consider $D_C(V_{S(n)})$ with vertices v_i and arcs a_i labelled so that a_i indicates $v_i v_{i+1}$ for every $i = 1, 2, 3, \ldots, k-1$. A (v_1, v_k)-path is a (v_1, v_k)-walk in $D_C(V_{S(n)})$ given by $v_1 a_1 v_2 a_2 v_3 \ldots a_{k-1} v_k$ such that all vertices are distinct. A *k-cycle* is Hamilton if it is a closed (v_1, v_k)-walk (i.e. $v_1 = v_k$) of length $k = n$ on the (v_1, v_{k-1})-path. A digraph is strongly connected (or strong) if for every pair $v_i, v_j \in V_{S(n)}$, $i, j = 1, 2, 3, \ldots, n$ and $i \neq j$, there exist both (v_i, v_j)-path and (v_j, v_i)-path [17]. Coevolutionary digraphs $D_C(V_{S(n)})$ are either reducible or irreducible. Any irreducible $D_C(V_{S(n)})$ on $n \geq 3$ vertices cannot be vertex-partitioned, is hamiltonian and strongly connected [12].

2.2. CMCs

We focus on population-one coevolutionary systems with processes described by random walks on digraphs. They also naturally represent learning algorithms based on self play [18].. A standard random walk on a labelled D_C starts at a random vertex, and in each subsequent step, jumps to one of the out-neighbours $N_D^+(u) = \{v \in V_S \backslash \{u\} : uv \in A_R\}$ of the current vertex u randomly with equal probability. They are modelled as a specific type of discrete time Markov chain $\mathbf{\Phi} = \{\Phi_t : t \in \mathbb{N}_0\}$, each Φ_t takes values from the countable state space $X = V_{S(n)}$. By exploiting time homogeneity and memory loss [19], we can construct the CMC $\mathbf{\Phi}$ as the random walk on $D_C(V_{S(n)})$ with initial distribution $\boldsymbol{\mu}$ over and Markov transition matrix $\boldsymbol{P}$ on $X = V_{S(n)}$ satisfying: (1) $\boldsymbol{\mu} = (\mu_x : x \in X)$, $0 \leq \mu_x \leq 1$ and $\sum_{x \in X} \mu_x = 1$, and (2) $\boldsymbol{P} = (\mathbb{P}(x, z) : x, z \in X)$, where every row is a distribution with $0 \leq \mathbb{P}(x, z) \leq 1$ and $\sum_{y \in X} \mathbb{P}(x, y) = 1$. The distribution describing $\mathbf{\Phi}$ can be obtained from $\boldsymbol{\mu}$ and $\boldsymbol{P}$.

There is an intimate connection between qualitative characterizations of global connectivity structures in D_C and quantitative characterizations of $\mathbf{\Phi}$ [12]. A random walk on a reducible D_C leads to an absorbing $\mathbf{\Phi}$. In the same manner, an irreducible $\mathbf{\Phi}$ operates on a strongly connected D_C. One can obtain the long-term limiting distribution of $\mathbf{\Phi}$ that satisfies the invariant property $\boldsymbol{\pi P} = \boldsymbol{\pi}$. $\boldsymbol{\pi} = (\pi_x : x \in X)$ on X is the stationary distribution of $\mathbf{\Phi}$. Each π_x describes the long-term fraction of time spent on x by $\mathbf{\Phi}$. For absorbing $\mathbf{\Phi}$, the probability mass is concentrated on the absorbing class with all other states having zero probabilities [20].

2.3. PageRank Authority for Coevolutionary Digraphs

We reformulate the PageRank authority [21] using digraph-theoretic notations to make explicit its link to digraph structures. Let $u, v \in V_S$, the in-neighbourhood of u is $N_D^-(u) = \{v \in V_S \backslash \{u\} : vu \in A_R\}$, and the out-degree (number of outgoing arcs) of v is $d_D^+(v) = |\{(u,v) \in A_R : v \in V_S \backslash \{u\}\}|$. The PageRank authority of u is $\varphi_u = \alpha + (1-\alpha) \sum_{v \in N_D^-(u)} \varphi_v / d_D^+(v)$ where $\alpha \in (0,1)$.

Let vertices in $D_C(V_{S(n)})$ be labelled as $v_1, v_2, v_3, \ldots, v_n$. The PageRank authorities are $\boldsymbol{\varphi} = [\varphi_1, \varphi_2, \varphi_3, \ldots, \varphi_n]$ (we write φ_{v_i} as the ith element φ_i of $\boldsymbol{\varphi}$), and given by $\boldsymbol{\varphi} = \alpha \mathbf{e} + (1-\alpha)\boldsymbol{\varphi}\mathbf{M}$, where $\mathbf{e} = [1, 1, 1, \ldots, 1]$ and the matrix $\mathbf{M} = (m_{ij} : i, j \in \{1, 2, 3, \ldots, n\})$ with $m_{ij} = \frac{1}{d_D^+(v_j)}$ if $j \to i$, $m_{ij} = 0$ if $j \nrightarrow i$. The influence of the digraph structure can be seen from non-zero entries of $\mathbf{M}$ that corresponds to the domination matrix associated with $D_C(V_{S(n)})$.

2.4. CMC and PageRank Connection

Actual PageRank computation requires reformulation as an eigensystem or an equivalent linear system if $\boldsymbol{\psi}$ is suitably normalized ($\boldsymbol{\psi}\mathbf{e}^T = 1$) [14]. PageRank is a Markov chain with a primitive probability transition matrix $\boldsymbol{P}_{\mathrm{PR}}$ (non-negative, irreducible matrix having one eigenvalue $r = \rho(\boldsymbol{P}_{\mathrm{PR}})$ on its spectral circle [22]). Similarly, setting the usual adjustment $\boldsymbol{P}_{\mathrm{PR}} = \alpha \mathbf{e}^T \mathbf{s} + (1-\alpha)\boldsymbol{P}$ where $\mathbf{s}$ is a general probability vector (e.g. uniform $\frac{1}{n}\mathbf{e}$) [23] (i.e. introducing restart with probability α) makes CMC irreducible and akin to a PageRank on $D_C(V_{S(n)})$. $\boldsymbol{P} = \mathbf{M}$ if standard random walk is used. $\boldsymbol{\psi}$ is the solution to $\boldsymbol{\psi}\boldsymbol{P}_{\mathrm{PR}} = \boldsymbol{\psi}$. This gives two natural interpretations to $\boldsymbol{\psi}$: (1) It indicates how coevolution measures the importance (authority) of individual $v \in V_{S(n)}$ and provides the means to rank them. (2) It gives long-term visitation probabilities over the digraph by coevolutionary search.

We will establish several theoretical results connecting CMC and PageRank. We focus on the personalized PageRank [24]. In the following, we provide guarantees on the existence of a PageRank vector associated with a coevolutionary system and a linear systems formulation that allows us to uncover several properties related to the coefficient matrix.

Lemma 1. *Let $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$ be a coevolutionary digraph. Let $\boldsymbol{P}$ be the probability transition matrix associated with a CMC operating on $D_C(V_{S(n)})$ and $\boldsymbol{Z} = (\mathbf{I}+\boldsymbol{P})/2$ its lazy version. The row vector $\mathbf{s}$ is the probability distribution over the set vertices $V_{S(n)}$. Given $\alpha \in (0,1)$ and $\beta = \frac{2\alpha}{1-\alpha}$, the personalized PageRank vector is the unique solution to the linear system defined as*

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) = \alpha\mathbf{s} + (1-\alpha)\boldsymbol{\psi}_\alpha(\mathbf{s})\boldsymbol{Z}. \tag{1}$$

and can be computed as

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) = \beta\mathbf{s}\big(\beta\mathbf{I} + (\mathbf{I}-\boldsymbol{P})\big)^{-1}. \tag{2}$$

Proof: Equation 1 is formulated in [25]. We can rewrite it as

$$\boldsymbol{\psi}_\alpha(\mathbf{s})\big(\beta\mathbf{I} + (\mathbf{I}-\boldsymbol{P})\big) = \beta\mathbf{s} \tag{3}$$

where $\beta > 0$ (see Appendix A). $\mathbf{W} = \beta\mathbf{I}+(\mathbf{I}-\boldsymbol{P})$ is a strictly dominant diagonal matrix and is invertible [26]. □

Lemma 2. *Let $\boldsymbol{\psi}_\alpha(\mathbf{s})\big(\beta\mathbf{I} + (\mathbf{I}-\boldsymbol{P})\big) = \beta\mathbf{s}$ be the formulation of the PageRank problem as a linear systems. The coefficient matrix $\mathbf{W} = \beta\mathbf{I} + (\mathbf{I}-\boldsymbol{P})$ has the following properties:*

1. $\mathbf{W}$ *is an M-matrix.*
2. $\mathbf{W}$ *is nonsingular.*
3. *The row sums of* $\mathbf{W}$ *are* β.
4. $||\mathbf{W}||_\infty = 2+\beta$.
5. $\mathbf{W}^{-1} \geq \mathbf{0}$.
6. *The row sums of* $\mathbf{W}^{-1}$ *are* $\frac{1}{\beta}$.
7. $||\mathbf{W}^{-1}||_\infty = \frac{1}{\beta}$.
8. $\kappa_\infty(\mathbf{W}) = \frac{2+\beta}{\beta} = \frac{1}{\alpha}$.

Proof: Let $\mathbf{R}^{n\times n}$ be the set of real, square $n \times n$ matrices. $\mathbf{A} \in \mathbf{R}^{n\times n}$ is an M-matrix with the form $\mathbf{A} = c\mathbf{I} - \mathbf{B}$, where $\mathbf{B} \geq \mathbf{0} = (b_{ij} \geq 0 : 1 \leq i, j \leq n)$ and $c \geq \rho(\mathbf{B})$ with $\rho(\cdot)$ denoting the spectral radius [27]. The main idea uses the fact that by definition, $\mathbf{W} = c\mathbf{I} - \mathbf{B}$ is an M-matrix with $c = 1 + \beta$ and $\mathbf{B} = \boldsymbol{P}$ a stochastic matrix. Various properties can be shown such as *inverse-positivity* (Theorem 2.3, [27]). See Appendix A for complete details. □

$\kappa_\infty(\mathbf{W})$ in Lemma 2 indicates the sensitivity of the solution to PageRank to perturbations in $\mathbf{W}$. It quantifies the extent $\mathbf{W}$ is ill-conditioned with respect to machine precision when direct computation based on Gaussian elimination is used. For a machine precision $\mathbf{u}$, this quantity is taken to be $\mathbf{u}\kappa_\infty(\mathbf{W})$ (Chapter 3, [28]). For example, $\mathbf{u}\kappa_\infty(\mathbf{W}) \leq 1$ would indicate a loss of a single decimal digit of precision for $\alpha = 0.1$. In practice, iterative methods are used for PageRank computation (see Theorem 2.2 [14] for error characterizations of these methods).

2.5. CMC and PageRank - Stationary Distributions

Given the connection between CMC and PageRank, the immediate issues are (1) characterization of introducing restart in random walks on D_C and (2) how changes to α impact the long-term limiting distribution of CMC. Our theoretical results address these issues qualitatively and quantitatively, for (1) and (2), respectively. We use recent results from spectral graph theory generalized to digraphs in [29]. Let $\boldsymbol{P}$ be associated with a CMC operating on irreducible $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$ with stationary vector $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3, \ldots, \pi_n)$ where $\boldsymbol{\pi P} = \boldsymbol{\pi}$. Let the diagonal matrix be $\mathbf{\Pi} = \mathrm{diag}(\pi_i)$. The normalized digraph Laplacian is $\tilde{\boldsymbol{\mathcal{L}}} = \mathbf{\Pi}^{\frac{1}{2}}(\mathbf{I} - \boldsymbol{P})\mathbf{\Pi}^{-\frac{1}{2}}$, with its individual element given by

$$\tilde{\mathcal{L}}_{ij} = \begin{cases} 1 - p_{ii} & \text{if } i = j \\ -\pi_i^{1/2} p_{ij} \pi_j^{1/2} & \text{if } (i,j) \in A \\ 0 & \text{otherwise.} \end{cases}$$

The Green's function for digraphs $\tilde{\boldsymbol{\mathcal{Z}}} = \tilde{\boldsymbol{\mathcal{L}}}^+$ is the Moore-Penrose pseudoinverse of $\tilde{\boldsymbol{\mathcal{L}}}$ with $\tilde{\boldsymbol{\mathcal{Z}}}\tilde{\boldsymbol{\mathcal{L}}} = \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}} = \mathbf{I} - \tilde{\boldsymbol{\mathcal{J}}}$, where $\tilde{\boldsymbol{\mathcal{J}}} = (\boldsymbol{\pi}^{1/2})^T \boldsymbol{\pi}^{1/2}$. The next theorem shows the relationship between PageRank and stationary vectors of CMC.

Theorem 3. *Let* $\boldsymbol{\pi}$ *be the stationary vector associated with a CMC operating on an irreducible* $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$*. The personalized PageRank vector is given by*

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) = \mathbf{s}\big(\mathbf{I} - \boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1}\boldsymbol{\Pi}^{\frac{1}{2}}\big). \tag{4}$$

Furthermore,

$$\lim_{\beta\to 0} \tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1} = \tilde{\boldsymbol{\mathcal{L}}}(\tilde{\boldsymbol{\mathcal{L}}})^{+} = \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}$$

where $\beta = \frac{2\alpha}{1-\alpha}$ *and* $\tilde{\boldsymbol{\mathcal{Z}}} = \tilde{\boldsymbol{\mathcal{L}}}^{+}$ *is the Moore-Penrose pseudoinverse of* $\tilde{\boldsymbol{\mathcal{L}}}$*. For small values of* α *(subsequently, small values of* β*), the PageRank vector is approximated by the stationary vector, i.e.,* $\boldsymbol{\psi}_\alpha(\mathbf{s}) \approx \boldsymbol{\pi}$ *with equality (*$\boldsymbol{\psi}_0(\mathbf{s}) = \boldsymbol{\pi}$*) for any* $\mathbf{s}$*.*

<u>Proof:</u> Appendix A gives the full proof while we summarize the main ideas here. Using various identities involving the normalized digraph Laplacian and the generalized identity for the inverse of a sum of matrices [30], we can rewrite Equation 2 so that

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) = \mathbf{s}\big(\mathbf{I} - \boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1}\boldsymbol{\Pi}^{\frac{1}{2}}\big)$$

can be calculated $(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1}$ is nonsingular (Theorem 1.2.17, [26]).

We can reformulate the linear system as

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) = \boldsymbol{\pi} + \beta(\mathbf{s} - \boldsymbol{\pi})\big(\beta\mathbf{I} + \mathbf{I} - \boldsymbol{P}\big)^{-1}, \tag{5}$$

which indicates that $\boldsymbol{\psi}_0(\mathbf{s}) = \boldsymbol{\pi}$ for any $\mathbf{s}$ as we let $\alpha \to 0 \Rightarrow \beta \to 0$.

We can directly calculate $\boldsymbol{\psi}_0(\mathbf{s})$. Let $\mathbf{N}_\beta = \beta(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1}$. We will apply results from (Theorem 1 and Lemma 1 [29]) involving identities relating the normalized digraph Laplacian and its Moore-Penrose pseudoinverse. As $\beta \to 0$, we have $\mathbf{N}_0 = \mathbf{N}_0\tilde{\boldsymbol{\mathcal{J}}}$. Given that $\tilde{\boldsymbol{\mathcal{J}}}$ is singular, the solutions to the equality $\mathbf{N}_0 = \mathbf{N}_0\tilde{\boldsymbol{\mathcal{J}}}$ are $\mathbf{N}_0 = \mathbf{0}$ and $\mathbf{N}_0 = \tilde{\boldsymbol{\mathcal{J}}}^k = \tilde{\boldsymbol{\mathcal{J}}}$, $k = 1, 2, 3, \ldots$. Then, $\tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1} = \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\big(\mathbf{I} - \beta(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1}\big)$ reduces to $\tilde{\boldsymbol{\mathcal{L}}}(\tilde{\boldsymbol{\mathcal{L}}})^{+} = \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}$, implying that

$$\lim_{\beta\to 0} \tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1} = \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}.$$

□

Theorem 3 shows a direct relationship between $\boldsymbol{\pi}$ and $\boldsymbol{\psi}_\alpha(\mathbf{s})$ as a result of introducing restart in CMC. Crucially, the connection between structures in D_C and quantities arising from random walks on D_C is made explicit through the digraph Laplacian. For small $\alpha > 0$, how the long-term fraction of time the random walker spends on each vertices that is redistributed is represented by the perturbed normalized digraph Laplacian $(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1}$. What is the difference between $\boldsymbol{\pi}$ and $\boldsymbol{\psi}_\alpha(\mathbf{s})$ due to restart with probability α? The following results answer this quantitatively.

Lemma 4. *Let a CMC operating on irreducible $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$ be associated with a probability transition matrix $\boldsymbol{P}$, stationary distribution $\boldsymbol{\pi}$, and the fundamental matrix Z. Correspondingly, the personalized PageRank on $D_C(V_{S(n)})$ is a CMC with probability transition matrix $\boldsymbol{P}_{PR}$ and stationary distribution $\boldsymbol{\psi}_\alpha(\mathbf{s})$. Then*

$$\begin{aligned}\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}) &= -\boldsymbol{\psi}_\alpha(\mathbf{s})\big(\mathbf{I} - \boldsymbol{P}\big)Z \\ &= \beta(\boldsymbol{\pi} - \mathbf{s})\big(\beta\mathbf{I} + (\mathbf{I} - \boldsymbol{P})\big)^{-1}.\end{aligned}$$

Proof: Applying results from Perturbation theory on finite Markov chains (Theorems 1 and 2, [31]), we obtain

$$\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}) = -\boldsymbol{\psi}_\alpha(\mathbf{s})(\mathbf{I} - \boldsymbol{P})Z.$$

We can show that the vector $\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})$ is invariant under multiplication of the matrix $(\mathbf{I} - \boldsymbol{P})Z$, i.e., $(\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}))(\mathbf{I} - \boldsymbol{P})Z = \boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})$. Furthermore, $\big((\mathbf{I} - \boldsymbol{P})Z\big)^k = (\mathbf{I} - \boldsymbol{P})Z$, $k = 1, 2, 3, \ldots$. From Equation 5, we obtain the following

$$\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}) = \beta(\boldsymbol{\pi} - \mathbf{s})\big(\beta\mathbf{I} + \mathbf{I} - \boldsymbol{P}\big)^{-1}.$$

Multiplying $(\mathbf{I} - \boldsymbol{P})Z$ on both sides of the equality above completes the proof (see Appendix A for details). □

Corollary 5. *For* $\boldsymbol{\pi}$ *and* $\boldsymbol{\psi}_\alpha(\mathbf{s})$ *associated with CMC on an irreducible* $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$*, the following inequality is given for restart probabilities* $\alpha_1 \leq \alpha_2$

$$||\boldsymbol{\pi} - \boldsymbol{\psi}_{\alpha_1}(\mathbf{s})|| \leq ||\boldsymbol{\pi} - \boldsymbol{\psi}_{\alpha_2}(\mathbf{s})|| \tag{6}$$

where $\alpha_1, \alpha_2 \in (0, 1)$.

Proof: Let $\mathbf{y} = \boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})$, which consists of entries that are differentiable functions of a real variable α [22]. We can show that $||\mathbf{y}(\alpha)||$ is monotonically increasing for α in $(0, 1)$. See Appendix A for complete details. □

Corollary 6. *Associated to each irreducible* $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$ *is CMC with stationary distribution* $\boldsymbol{\psi}_\alpha(\mathbf{s})$ *for* $\alpha \in (0, 1)$ *and* $\boldsymbol{\pi}$ *with the following perturbation bound*

$$||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty \leq ||\boldsymbol{\pi} - \mathbf{s}||_\infty \tag{7}$$

with equality when $\alpha = 1$.

Proof: Since $||(\beta\mathbf{I} + \mathbf{I} - \boldsymbol{P})^{-1}||_\infty = \frac{1}{\beta}$ from Lemma 2, we obtain $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty \leq \beta||\boldsymbol{\pi} - \mathbf{s}||_\infty \, ||(\beta\mathbf{I} + \mathbf{I} - \boldsymbol{P})^{-1}||_\infty = ||\boldsymbol{\pi} - \mathbf{s}||_\infty$. See Appendix A for details. □

In this paper, we adopt a centrality measure approach [14] and take $\mathbf{s}$ to be uniform for a global digraph analysis. For suitable norms, $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||$ is monotonic with α (Corollary 5). $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ is upper-bounded by $||\boldsymbol{\pi} - \mathbf{s}||_\infty$ (Corollary 6). A general upper-bound of $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1$ can be obtained for probability vectors associated with random walks on labelled (isomorphic) tournaments [32, 33], taking into account their unilateral and directional duals [34, 17] (see Appendix A)

$$||\boldsymbol{\pi}_1 - \boldsymbol{\pi}_2||_1 \leq 2\Big(1 - \frac{1}{n}\Big). \tag{8}$$

Although characterizing $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1$ against changes in α is more useful, tight bounds are difficult to obtain (e.g. $\frac{2}{\alpha}$) [35, 23, 21]. We combine a coupling approach [35] with digraph-theoretic arguments to improve the bound to $\frac{2}{1-\alpha}$ for irreducible D_C (see Lemma 5, Appendix A).

3. Computational Results

We present results from controlled empirical studies on CMCs operating on coevolutionary tournaments $T_C(V_{S(n)})$ having specific structures [32, 17, 12]. They are selected based on known internal cycle structures that let us test current understanding of those structures and evaluate their impact on the visitation probabilities of vertices by the coevolutionary random walkers using quantitative measures we develop earlier. We also compute differences on PageRank orderings [36] that we describe in detail later. We adapt the standard Power Method approach [23] to compute the PageRank vectors $\boldsymbol{\psi}_\alpha(\mathbf{s})$, using the lazy version of the CMCs with uniform (teleportation vector) $\mathbf{s}$.

3.1. Generating Coevolutionary Tournaments

The set of coevolutionary digraphs for the experiments include irreducible, reducible, and random $T_C(V_{S(n)})$. All irreducible $T_C(V_{S(n)})$ are pancyclic [12]. The vertex-pancyclic $T_C(V_{S(n)})$ with the least number of 3-cycles [37] can be obtained from the transitive tournaments of order n indexed by its score sequences and then just reversing the arc v_1v_n to v_nv_1. It has a single transitive subtournament induced from a maximal, disjoint vertex partition of $n-2$ vertices. We will refer to these tournaments as *pancyclic (maximal transitive subtournament)*. Other vertex-pancyclic (or simply pancyclic) $T_C(V_{S(n)})$ that we use are those where we further reverse the arc v_2v_{n-1} to $v_{n-1}v_2$.

For reducible cases, we use two approaches to generate reducible $T_C(V_{S(n)})$ with various degrees of cycle structures in a controlled manner. The first approach exploits known digraph-theoretic structures. Every reducible $T_C(V_{S(n)}) \in \mathcal{T}_C(V_{S(n)})n \geq 2$ has a strong decomposition $V\big(T^{(1)}\big) \cup V\big(T^{(2)}\big) \cup V\big(T^{(3)}\big) \cup \ldots \cup V\big(T^{(l)}\big)$ with a unique ordering $T^{(1)}, T^{(2)}, T^{(3)}, \ldots, T^{(l)}$ whereby $T^{(i)} \mapsto T^{(j)}$ when $i < j$ for $i, j = 1, 2, 3, \ldots, l$. $T^{(1)}$ $\big(T^{(l)}\big)$ is the initial (terminal) strong component [12]. We generate reducible $T_C(V_{S(n)})$ with odd number of components so that each odd-numbered-ith component is a single vertex and even-numbered-ith component consists of a strong component of known structures, e.g., vertex-pancyclic (maximal transitive subtournament) and regular [17].

The second approach uses a network-growth-based methodology via preferential attachment [12, 38]. A hierarchy of $T_C(V_{S(n)})$ where a control parameter is used to move between two complexity extremes – irreducible $T_C(V_{S(n)})$ and reducible $T_C(V_{S(n)})$ with prominent transitive structures – can be generated. We use the same settings in [12]: an initial transitive T_C of order ten is used as a seed and attachment probabilities $\mathbb{P}_i^- = \big(1 + \exp(-\gamma(x_i - \overline{x})\big)^{-1}$ of nodes v_i are computed with $\{x_i\}_1^n$ being the ranks of v_i sorted according to the score sequences, $\overline{x} = n/2$, and γ that plays the role of the inverse temperature. We control the generator to produce reducible T_C of order $n+1$ by introducing at the final iteration a single dominant node v_{n+1}.

Finally, random $T_C(V_{S(n)})$ is generated by orienting each edge $\{v_i, v_j\} \in E$ of its underlying, complete graph randomly with equal probability for each direction. However, a random $T_C(V_{S(n)})$ is likely to be irreducible (e.g. more than half the chance for a generated random tournament of order five [32]). Irreducibility test on random $T_C(V_{S(n)})$ can be done using their score sequences following a well-known result of Moser and Harary (1966) [17].

3.2. Results For $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||$

We evaluate both max and one norms and study the impact of the restart probability α on $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||$ as the structure and size of $T_C(V_{S(n)})$ are varied in a controlled manner. Fig. 1 shows the results for $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$. For reducible $T_C(V_{S(n)})$, they all have one absorbing state v_n so that $\boldsymbol{\pi} = (0, 0, 0, \ldots, 1)$. Following Theorem 3, $\boldsymbol{\psi}_\alpha(\mathbf{s})$ is a redistribution of $\boldsymbol{\pi}$ and amounts to leakages of the probability mass concentrated on v_n. Essentially, the max norm takes the value $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty = |\pi_{v_n} - \psi_{v_n}|$. Since there are leakages to a maximal subset $V_{S(n)} \backslash \{v_n\}$ of $n-1$ vertices, $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ for reducible cases would be larger in comparison with the irreducible cases where there are no leakages.

$||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ becomes larger and gets closer to the upper bound $(1 - \frac{1}{n})$ (bold line in Fig. 1) as a result of increasing number of 3-cycles that leads to fewer transitive components in the reducible $T_C(V_{S(n)})$s. This is achieved by setting the dominated strong component parts with specific cycle structures

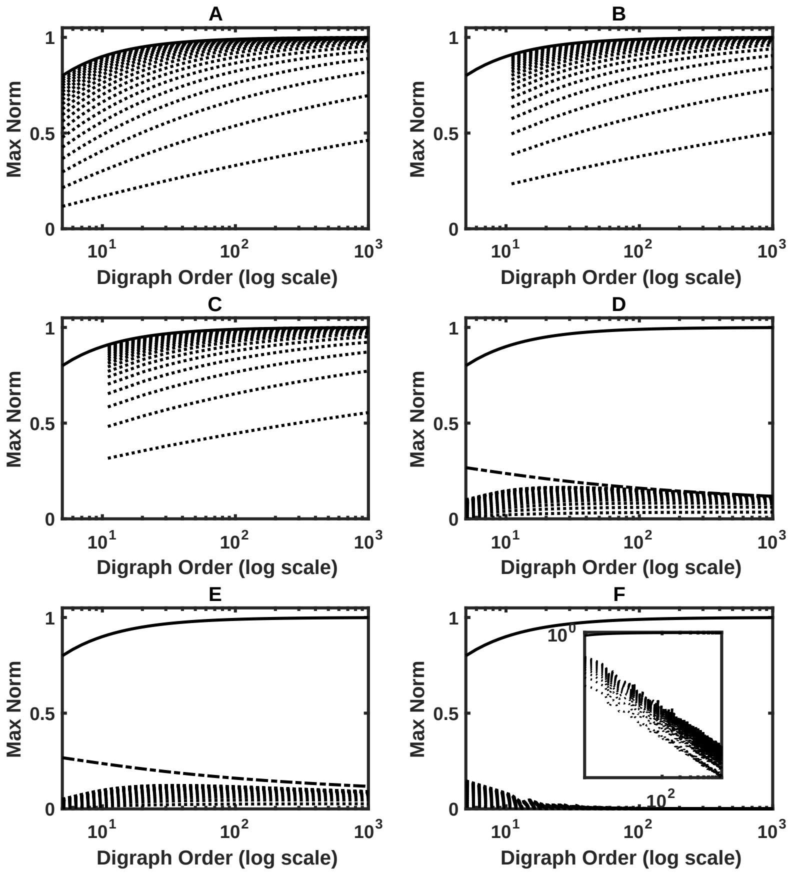

Figure 1: $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ of CMCs operating on tournaments that are (A) transitive, (B) with pancyclic (maximal transitive subtournament) components of order 9, (C) with regular components of order 9, (D) pancyclic (maximal transitive subtournament), (E) pancyclic, and (F) random tournaments of odd order $n \in [5, 999]$. All plots are semilog except the inset of F, which is log-log. Dotted lines indicate $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ for $\alpha \in [0.05, 0.95]$ in steps of 0.05. Upper bounds: $(1 - \frac{1}{n})$ (bold lines) and $(\frac{a_1}{a_2 + H_{m+1}} - \frac{1}{m+2})$ (dash-dot lines in D-E).

from pancyclic with the least number of 3-cycles in B to regular with the most number of 3-cycles (corollary of Theorem 4, [32]) in C.

For irreducible cases, both stationary and PageRank vectors will be nonnegative (all entries > 0) so that $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ are substantially lower than $(1 - \frac{1}{n})$.

Intuition suggests that as the number of 3-cycles increases, i.e., towards tournaments with regular cycle structures that have maximum number of 3-cycles and uniform $\boldsymbol{\pi}$ [12], $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ will decrease. This can be observed by comparing results in D for pancyclic $T_C(V_{S(n)})$ with the least number of 3-cycles and results in E and particularly in F where random $T_C(V_{S(n)})$ become increasingly more regular-like as the number of vertices increases (inset of F).

Computational results indicate that $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ is monotonic with α (see the higher-valued dotted lines in Fig. 1 representing results where larger α values are used for reducible and irreducible cases). $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ gets closer to the upper bound $||\boldsymbol{\pi} - \mathbf{s}||_\infty \leq ||\boldsymbol{\pi}_{\text{tran}} - \mathbf{s}||_\infty = (1 - \frac{1}{n})$ for the reducible A-C cases and $||\boldsymbol{\pi} - \mathbf{s}||_\infty$ for the irreducible D-E cases. Although this property was shown explicitly only for irreducible CMCs in Corollary 5, one can argue from the point of leakages and setting a uniform $\mathbf{s}$ that $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ is monotonic for reducible CMCs too.

Sharp bounds can be obtained for CMCs operating on $T_C(V_{S(n)})$ with specific structures in D that consist of pancyclic (maximal transitive subtournament). We note that $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty = |\pi_{v_1} - \psi_{v_1}| \leq ||\boldsymbol{\pi} - \mathbf{s}||_\infty = |\pi_{v_1} - \frac{1}{n}|$. This coincides with our intuitive understanding how this structure affects on PageRank calculation. Redistribution affects vertex v_1, which shares most of the probability mass with v_n. As in [21], the PageRank authority of v_1 is obtained from a single link $v_n \to v_1$ even though v_n has the highest authority.

We can calculate this bound directly using digraph-theoretic arguments (see notes after Corollary 8, Appendix A). We have these identities: $H_n = \sum_{k=1}^{n} \frac{1}{k}$ (nth Harmonic number), $\sum_{i=1}^{n} H_i = (n+1)(H_{n+1} - 1)$ [39], and $\mathbb{E}(\eta_{V_{S(n-2)}}) = \frac{1}{n-2} \sum_{i=1}^{n-2} H_i$. We apply these identities with $m = n-2, n \geq 3$ on Equation A.9 in Appendix A to obtain

$$\begin{aligned}\pi_{v_1} &= \frac{1}{2}\left(1 - \frac{\frac{1}{m}\sum_{i=1}^{m} H_i}{2 + \frac{1}{m}\sum_{i=1}^{m} H_i}\right) \\ &= \frac{1}{2}\left(1 - \frac{\frac{1}{m}(m+1)(H_{m+1} - 1)}{2 + \frac{1}{m}(m+1)(H_{m+1} - 1)}\right)\end{aligned}$$

$$= \frac{a_1}{a_2 + H_{m+1}}$$

where $a_1 = \frac{m}{m+1}$ and $a_2 = \frac{m-1}{m+1}$. Then,

$$||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty \leq \Big(\frac{a_1}{a_2 + H_{m+1}} - \frac{1}{m+2}\Big).$$

For these tournaments with large but finite number of vertices $n = m + 2$, the right-hand side of the inequality is dominated by the reciprocal of H_{m+1} (since H_{m+1} grows logarithmically) and is bounded away from zero [39]. Equality is obtained only for $m = 1$ where the right-hand side of the inequality is zero. This corresponds to the isomorphic pancyclic tournament of $n = 3$ vertices, which is regular and has $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})|| = 0$.

Fig. 2 shows the results for $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1$ for the same set of experiments. The upper-bound $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1 \leq 2(1 - \frac{1}{n})$ is plotted as a bold line while $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1 \leq \frac{2}{1-\alpha}\Big(1 - \big(\pi_{v_n} - \frac{1}{n}\big)\Big)$ is plotted as dash-dot line in Fig. 2. For reducible A-C cases with one absorbing state v_n, we can verify numerically that $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1 = 2||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$. Any leakage that is measured by the max norm would be redistributed over all other states. The one norm cumulatively adds these leakages in addition to that given by the max norm for v_n. Crucially, this leakage relationship extends to that between the digraph-theoretic-derived upper-bounds since $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1 \leq 2(1 - \frac{1}{n})$ and $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty \leq (1 - \frac{1}{n})$.

We now turn our attention to experiments involving reducible $T_C(V_{S(n)})$ generated from the network-growth-based methodology and consider $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ only. Fig. 3 shows results where the temperature $1/\gamma$ for the system is increased from A to D. At a lower temperature (A), random orientations of edges generated as a result of introducing a new node at each iteration of the method are biased towards existing nodes with higher scores. This leads to more prominent transitive structures in $T_C(V_{S(n)})$, detected by computing its Landau's index ν [12] that gives the number of strong components in $T_C(V_{S(n)})$. ν ranges at $[8, 49]$, $[8, 39]$, $[7, 47]$ for experiments with the final generated $T_C(V_{S(n)})$ at $n = 100$, 500, and 1000, respectively. At higher temperatures, $T_C(V_{S(n)})$ generated in B-D have two strong components only (a single dominant node and a strong

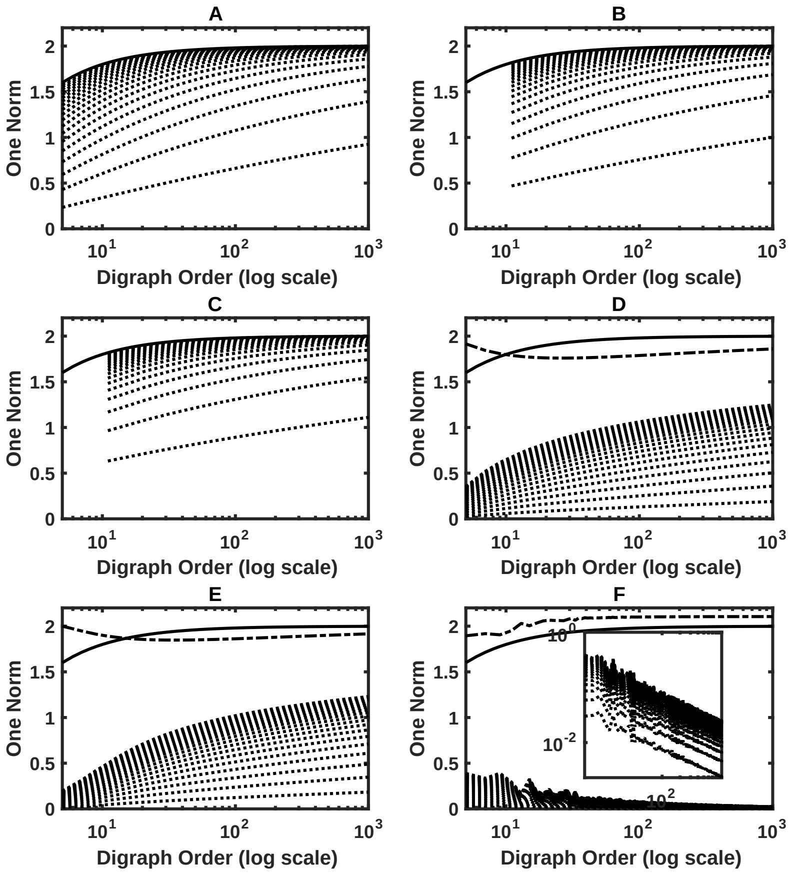

Figure 2: $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1$ of CMCs operating on tournaments that are (A) transitive, (B) with pancyclic (maximal transitive subtournament) components of order 9, (C) with regular components of order 9, (D) pancyclic (maximal transitive subtournament), (E) pancyclic, and (F) random tournaments of odd order $n \in [5, 999]$. All plots are semilog except the inset of F, which is log-log. Dotted lines indicate $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1$ for $\alpha \in [0.05, 0.95]$ in steps of 0.05. Upper bounds: $2(1 - \frac{1}{n})$ (bold lines) and $\frac{2}{1-\alpha}\left(1 - \left(\pi_{v_n} - \frac{1}{n}\right)\right)$ at $\alpha = 0.05$ setting (dash-dot lines in D-F).

dominated component of $n-1$ nodes). A higher $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ value is a result of the larger-sized, strong dominated component drawing out more probability mass away from the absorbing dominant node (more leakages).

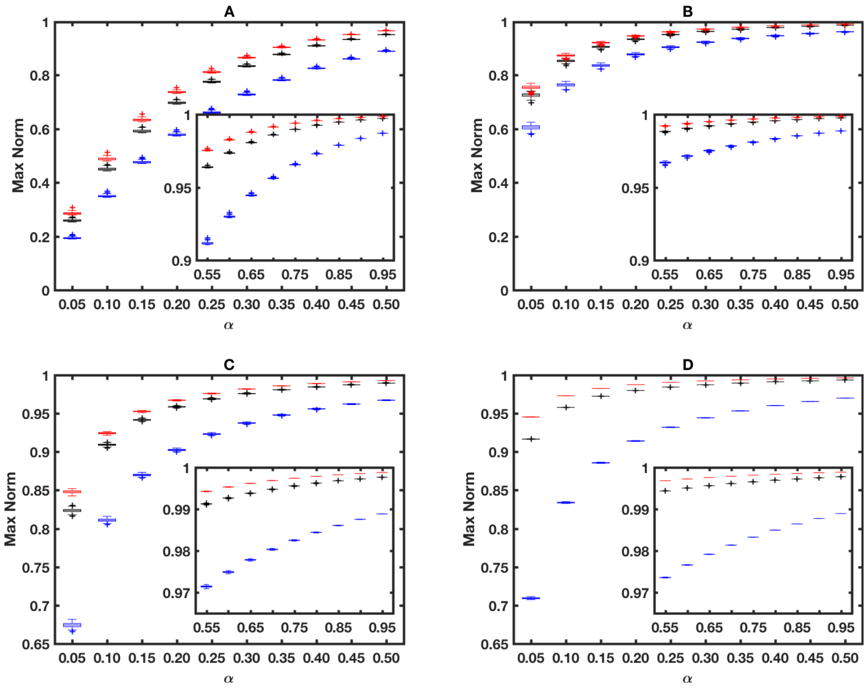

Figure 3: $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty$ of CMCs operating on reducible tournaments that are generated from a network-growth-based methodology with γ set to (A) 2.0, (B) 0.1, (C) 0.05, and (D) 0.01. Box and line plots in blue, black, and red represents results where the final generated tournaments are of order 100, 500, and 1000, respectively.

3.3. Results For PageRanking Coevolutionary Tournaments

The PageRank vector $\boldsymbol{\psi}_\alpha(\mathbf{s})$ gives the visitation probabilities ψ_{v_i} that ranks the importance (performance) of vertices $v_i \in V_{S(n)}$ representative of the underlying structure of their pairwise relations in the coevolutionary digraph $D_C(V_{S(n)})$. We PageRank (index) the importance of these vertices in the same manner as their score sequence for consistency. Let $\tau : \{\psi_{v_i}\}_i^n \to \mathbb{Z}_k^n$ be the PageRanks of the vertices $v_i \in V_{S(n)}$ of $D_C(V_{S(n)})$, with visitation probabilities $\{\psi_{v_i}\}_i^n$ and $\mathbb{Z}_k^n = \{1, 2, 3, \ldots, n\}$. Vertices are *PageRanked* in the ascending order from the least important $\tau(v_1) = 1$ to the most important $\tau(v_n) = n$. Ties are handled in the usual manner through fractional (average) ranks [40]. In our experiments, we generate tournaments and strong components of odd orders so that tied ranks can be assigned with the median (integer) value.

For transitive $T_C(V_{S(n)})$, results show that the ranking of vertices from τ is the same as that of score sequences $\left(d_T^-(v_i)\right)_i^n$ even for high α setting. How about irreducible $T_C(V_{S(n)})$ with prominent transitive structures? In the case of the pancyclic (maximal transitive subtournament) $T_C(V_{S(n)})$, visitation probabilities are sorted as $\psi_{v_2} < \psi_{v_3} < \psi_{v_4} < \cdots < \psi_{v_{n-1}} < \psi_{v_1} < \psi_{v_n}$, for reasonably low α settings. Unlike the situation with the stationary probability vector where the two highest ranks are tied since $\pi_{v_1} = \pi_{v_n}$, PageRank is able to distinguish between v_1 and v_n.

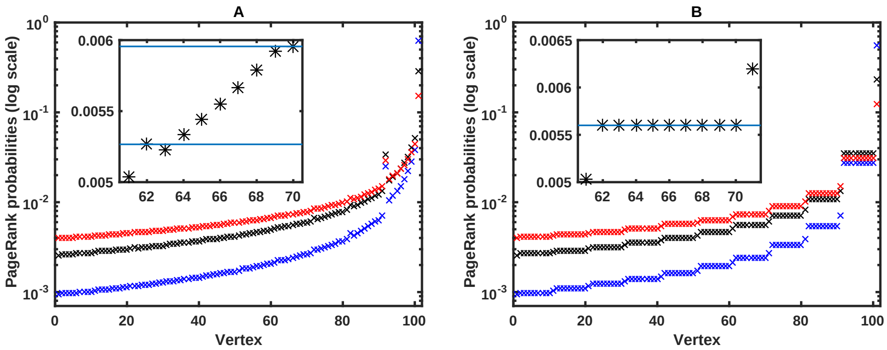

Figure 4: Visitation probabilities (PageRank vector) of the CMC with restart on the vertices of reducible tournaments of order 101 (A) with pancyclic (maximal transitive subtournament) components of order 9 and (B) with regular components of order 9. Plots with blue, black, and red crosses are for the visitation probabilities where α is set to 0.05, 0.15 and 0.25, respectively. Insets show the visitation probabilities where $\alpha = 0.25$ for one of the strong components. Vertices are sorted according to their score sequences.

If such a structure is embedded within a reducible $T_C(V_{S(n)})$ (as one may encounter in real-world problems with such complex structures), PageRank can be used to uncover this strong component structure for small-sized $T_C(V_{S(n)})$ where the issue of scale does not arise although the analysis can be more nuanced. For example, consider the case whereby the strong component is pancyclic (maximal transitive subtournament). Nodes v_{62} and v_{63} (inset of Fig. 4A) have the same out-degree globally ($d_T^-(v_{62}) = d_T^-(v_{63})$) and within the strong component, the least number of just one outgoing link locally. However, v_{62} has a higher PageRank than v_{63}. Together both information would indicate that they are part of a strong component that dominates all other vertices $\{v_1, v_2, v_3, \ldots, v_{61}\}$. If the strong component is regular, PageRanks are the same for those vertices (inset of Fig. 4B). Regardless, our setting (uniform $\mathbf{s}$) is to provide a global view [14]. For reducible $T_C(V_{S(n)})$, there would be a general increase in visitation probabilities ψ_{v_i} for vertices with higher scores and that vertices in the absorbing class have significantly larger values (in Fig. 4, both tournaments are reducible with a single dominant vertex).

3.4. Results For Differences in Rank Aggregation

It is known that setting α appropriately can improve the convergence rate of the Power Method (Theorem 5.1 in [23]) to compute PageRanks at the expense of increasing perturbations on PageRank authorities and subsequently the actual PageRank orderings [36]. We study the impact of α on PageRanks by computing the differences between PageRanks at baseline $\alpha = 0.05$ setting and those obtained from using higher α settings. There are various approaches to calculate this difference through rank aggregation. We use a form of the Spearman footrule distance [41].

For $T_C(V_{S(n)})$ of odd orders n, we can calculate this distance

$$d_{\mathrm{spear}}(\tau_1, \tau_2) = \frac{4}{n^2 - 1} \sum_{i=1}^{n} |\tau_1(v_i) - \tau_2(v_i)| \tag{9}$$

where the normalizing constant is given by the maximum of the sum of absolute differences $\max d(\tau_1, \tau_2) = \frac{n^2-1}{4}$. This can be obtained in the same manner as the Spearman's rank correlation [40] by calculating the distance obtained from completely opposed rankings of $1, 2, 3, \ldots, n$ and $n, n-1, n-2, \ldots, 1$ that also take into account directional duals of labelled tournaments [34]. The maximum distance is the same as in the case of the distance between that of transitive and regular tournaments. We also compute the average rank difference

$$d_{\mathrm{avg}}(\tau_1, \tau_2) = \frac{1}{n} \sum_{i=1}^{n} |\tau_1(v_i) - \tau_2(v_i)|. \tag{10}$$

$d_{\text{avg}}(\tau_1, \tau_2)$ would be a useful comparison with $d_{\text{spear}}(\tau_1, \tau_2)$ for experiments where the n^2 factor could mask the effect of small rank differences for certain large tournaments.

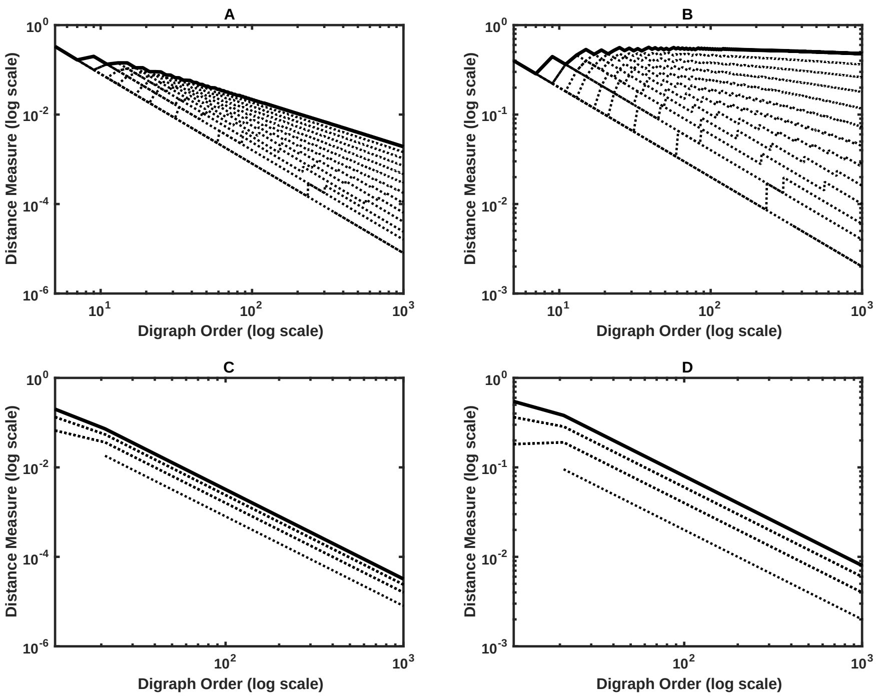

Figure 5: Differences of PageRanks for coevolutionary tournaments that are irreducible having (A)-(B) pancyclic (maximal transitive subtournament) structures and reducible having (C)-(D) pancyclic (maximal transitive subtournament) components of order 9. Plots in (A) and (C) are for $d_{\text{spear}}(\tau_1, \tau_2)$ and (B) and (D) are for $d_{\text{avg}}(\tau_1, \tau_2)$. Dotted lines show difference from baseline PageRanks obtained with $\alpha = 0.05$ to PageRanks obtained from setting higher $\alpha \in [0.10, 0.90]$. Bold lines are the case where $\alpha = 0.95$.

In general, higher α settings would lead to more perturbations to PageRanks measured by larger $d(\tau_1, \tau_2)$ values (e.g. the bold lines in Fig. 5). However, increasing the size of the generated $T_C(V_{S(n)})$ leads to lower $d(\tau_1, \tau_2)$ values, indicating that changes to PageRanks are localized to certain vertices. For example, this is usually v_1 for pancyclic (maximal transitive subtournament) $T_C(V_{S(n)})$ in A-B. Beyond specific structures in $T_C(V_{S(n)})$, our results indicate that the presence of prominent transitive structures (A-B) and those that forces reducibility (C-D) would localize perturbations of PageRanks to a small number of vertices, leading to lower $d(\tau_1, \tau_2)$ values as $T_C(V_{S(n)})$ increases in size.

4. Conclusion

One important issue dealing with problems in competitive settings is how problem structures can affect the performance of solutions discovered by iterative means. Under some assumptions, coevolutionary processes can be represented as Markov chain models operating on digraphs that capture preference relations between competing solutions to those problems. As a consequence, we are able to establish a direct and formal connection between the Markov chain model of coevolutionary processes and PageRank operating on such digraphs. This lets us develop a principled approach for quantitative characterization of coevolutionary problems regardless of the underlying structures in their digraph representations. This PageRank characterization of a coevolutionary digraph measures and ranks of the performance of individual solutions.

Our theoretical support involves interdisciplinary studies in large-scale coevolutionary systems, Markov models of PageRank and digraph theory. We provide guarantees of the existence of PageRank for any coevolutionary system. We prove the PageRank vector to be a redistribution of the stationary vector associated with CMCs operating on coevolutionary digraphs. Changes as a result of introducing restarts with different probability α can be quantified precisely. Our theoretical bounds are loose to cover for digraphs of any cycle structure but there are cases where qualitative knowledge of digraph structures can be exploited to obtain sharp bounds. We consider a restricted class of population-one coevolutionary systems, which covers other learning algorithms involving self play. The advantage is the coevolutionary digraph captures all underlying structures induced by the pairwise preferences over solutions for the problem under consideration. Specific reducible digraphs would cover evolutionary cases as in recent studies that apply PageRank to evolutionary systems [42].

Empirical studies have demonstrated how PageRank authorities characterize coevolutionary tournaments of various degree of complexity and known structures. The challenge remains to deal with large, real-world coevolutionary systems. New methods that can discover and exploit digraph structures (locally and globally) would be needed to obtain the relevant characterizations of these large-scale and complex systems. Unlike other problem domains with sparse $\boldsymbol{P}$ [14] and for which there exists methods to speed up computation [23], there are at least half of nonzero entries in $\boldsymbol{P}$ for coevolutionary systems by definition. There are scaling and computational issues if a global analysis perspective is adopted but this can be mitigated if some form of structures can be exploited (grouping vertices that represent a type of strategies of similar behaviors).

Acknowledgment

This work was supported by a H2020-MSCA-IF-2014 grant (No. 657027) on "COEVOLFRAMEWORK – Unified Framework for the Analysis of Coevolutionary Systems".

Appendix A. Complete Proofs for Theoretical Results Listed in Main Texts

Lemma 1. *Let $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$ be a coevolutionary digraph. Let $\boldsymbol{P}$ be the probability transition matrix associated with a CMC operating on $D_C(V_{S(n)})$ and $\boldsymbol{Z} = (\mathbf{I} + \boldsymbol{P})/2$ its lazy version. The row vector $\mathbf{s}$ is the probability distribution over the set vertices $V_{S(n)}$. Given $\alpha \in (0, 1)$ and $\beta = \frac{2\alpha}{1-\alpha}$, the personalized PageRank vector is the unique solution to the linear system defined as*

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) = \alpha\mathbf{s} + (1-\alpha)\boldsymbol{\psi}_\alpha(\mathbf{s})\boldsymbol{Z}. \tag{A.1}$$

and can be computed as

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) = \beta\mathbf{s}\big(\beta\mathbf{I} + (\mathbf{I} - \boldsymbol{P})\big)^{-1}. \tag{A.2}$$

Proof: We can rewrite Equation A.1 as follows

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) = \alpha\mathbf{s} + (1-\alpha)\boldsymbol{\psi}_\alpha(\mathbf{s})\boldsymbol{Z}$$

$$\boldsymbol{\psi}_\alpha(\mathbf{s})\big(\mathbf{I} + \frac{(1-\alpha)}{2\alpha}(\mathbf{I}-\boldsymbol{P})\big) = \mathbf{s}$$

so that now it can be expressed in the usual matrix form for linear systems

$$\boldsymbol{\psi}_\alpha(\mathbf{s})\big(\beta\mathbf{I} + (\mathbf{I}-\boldsymbol{P})\big) = \beta\mathbf{s} \tag{A.3}$$

where $\beta = \frac{2\alpha}{1-\alpha} > 0$.

$\beta\mathbf{I} + (\mathbf{I}-\boldsymbol{P})$ is a *strictly dominant diagonal matrix* and so is nonsingular (invertible) [26]. This can be proven as follows. Let $\mathbf{W} = \beta\mathbf{I} + (\mathbf{I}-\boldsymbol{P}) = (1+\beta)\mathbf{I}-\boldsymbol{P}$. Since $\boldsymbol{P}$ is a row stochastic matrix, then for each row i, $\sum_{j\neq i}|w_{ij}| \leq 1$. But $|w_{ii}| = 1+\beta > \sum_{j\neq i}|w_{ij}|$ for $\beta > 0$. □

Lemma 2. *Let $\boldsymbol{\psi}_\alpha(\mathbf{s})\big(\beta\mathbf{I} + (\mathbf{I}-\boldsymbol{P})\big) = \beta\mathbf{s}$ be the formulation of the PageRank problem as a linear systems. The coefficient matrix $\mathbf{W} = \beta\mathbf{I} + (\mathbf{I}-\boldsymbol{P})$ has the following properties:*

1. $\mathbf{W}$ *is an M-matrix.*
2. $\mathbf{W}$ *is nonsingular.*
3. *The row sums of* $\mathbf{W}$ *are* β.
4. $||\mathbf{W}||_\infty = 2+\beta$.
5. $\mathbf{W}^{-1} \geq \mathbf{0}$.
6. *The row sums of* $\mathbf{W}^{-1}$ *are* $\frac{1}{\beta}$.
7. $||\mathbf{W}^{-1}||_\infty = \frac{1}{\beta}$.
8. $\kappa_\infty(\mathbf{W}) = \frac{2+\beta}{\beta} = \frac{1}{\alpha}$.

Proof: The properties of $\mathbf{W}$ and their proofs follow a similar exposition in [23]. Let $\mathbf{R}^{n\times n}$ be the set of real, square $n \times n$ matrices. $\mathbf{A} \in \mathbf{R}^{n\times n}$ is an M-matrix if it can be written in the form $\mathbf{A} = c\mathbf{I} - \mathbf{B}$, where $\mathbf{B} \geq \mathbf{0}$ (i.e., $\mathbf{B} = (b_{ij} \geq 0 : 1 \leq i,j \leq n)$) and $c \geq \rho(\mathbf{B})$ with $\rho(\cdot)$ denoting the spectral radius (Chapter 6, [27]). Then, we have the following.

1. $\mathbf{W}$ is an M-matrix: $\mathbf{W} = \beta\mathbf{I} + (\mathbf{I}-\boldsymbol{P}) = c\mathbf{I} - \mathbf{B}$ where $c = 1+\beta$ and $\mathbf{B} = \boldsymbol{P}$. Applying the Perron-Frobenius Theorem to the stochastic matrix $\boldsymbol{P}$, $\rho(\boldsymbol{P}) = ||\boldsymbol{P}||_\infty = 1$ (Chapter 8, [22]). Obviously, $c = 1+\beta > \rho(\boldsymbol{P}) = 1$ for $\beta > 0$. As such, $\mathbf{W}$ is an M-matrix.

2. $\mathbf{W}$ is nonsingular: This was shown earlier in the proof for Lemma 1.
3. The row sums of $\mathbf{W}$ are β: Note that for each row i, $w_{ii} - \sum_{j \neq i} w_{ij} = (1+\beta) - 1 = \beta$. So, $\mathbf{W}\mathbf{e}^T = \beta\mathbf{e}^T$ where $\mathbf{e}^T$ is the column vector of ones.
4. $||\mathbf{W}||_\infty = 2 + \beta$: Applying the definition of ∞-norm [22], $||\mathbf{W}||_\infty = \max_i \sum_j |w_{ij}| = 2 + \beta$.
5. Since $\mathbf{W}$ is an M-matrix, then $\mathbf{W}^{-1} \geq \mathbf{0}$: This is simply direct application of (Theorem 2.3, Chapter 6, [27]). Since $\mathbf{W}$ is a nonsingular M-matrix, it is *inverse-positive*, i.e., $\mathbf{W}^{-1}$ exists and $\mathbf{W}^{-1} \geq \mathbf{0}$.
6. The row sums of $\mathbf{W}^{-1}$ are $\frac{1}{\beta}$: We note

$$\begin{aligned} \mathbf{W}\mathbf{e}^T &= \beta\mathbf{e}^T \\ \frac{1}{\beta}\mathbf{e}^T &= \mathbf{W}^{-1}\mathbf{e}^T. \end{aligned}$$

7. $||\mathbf{W}^{-1}||_\infty = \frac{1}{\beta}$: Since $\mathbf{W}^{-1} \geq \mathbf{0}$ and $\mathbf{W}^{-1}\mathbf{e}^T = \frac{1}{\beta}\mathbf{e}^T$, then $||\mathbf{W}^{-1}||_\infty = \frac{1}{\beta}$.
8. $\kappa_\infty(\mathbf{W}) = \frac{2+\beta}{\beta} = \frac{1}{\alpha}$: The condition number κ for a nonsingular $\mathbf{A}$ is $||\mathbf{A}||\,||\mathbf{A}^{-1}||$ [22]. Applying results in (4) and (7) with $\beta = \frac{2\alpha}{1-\alpha}$, we have

$$\begin{aligned} \kappa_\infty(\mathbf{W}) &= ||\mathbf{W}||_\infty\,||\mathbf{W}^{-1}||_\infty \\ &= \frac{2+\beta}{\beta} \\ &= \frac{1}{\alpha}. \end{aligned}$$

□

Theorem 3. *Let $\boldsymbol{\pi}$ be the stationary vector associated with a CMC operating on an irreducible $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$. The personalized PageRank vector is given by*

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) = \mathbf{s}\big(\mathbf{I} - \mathbf{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1}\mathbf{\Pi}^{\frac{1}{2}}\big). \tag{A.4}$$

Furthermore,

$$\lim_{\beta \to 0} \tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1} = \tilde{\boldsymbol{\mathcal{L}}}(\tilde{\boldsymbol{\mathcal{L}}})^{+} = \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}$$

where $\beta = \frac{2\alpha}{1-\alpha}$ and $\tilde{\boldsymbol{\mathcal{Z}}} = \tilde{\boldsymbol{\mathcal{L}}}^{+}$ is the Moore-Penrose pseudoinverse of $\tilde{\boldsymbol{\mathcal{L}}}$. For small values of α (subsequently, small values of β), the PageRank vector is *approximated by the stationary vector, i.e.,* $\boldsymbol{\psi}_\alpha(\mathbf{s}) \approx \boldsymbol{\pi}$ *with equality (*$\boldsymbol{\psi}_0(\mathbf{s}) = \boldsymbol{\pi}$*) for any* $\mathbf{s}$*.*

Proof: Since $\tilde{\boldsymbol{\mathcal{L}}} = \mathbf{\Pi}^{\frac{1}{2}}(\mathbf{I} - \boldsymbol{P})\mathbf{\Pi}^{-\frac{1}{2}}$ can be rewritten as $(\mathbf{I} - \boldsymbol{P}) = \mathbf{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\mathbf{\Pi}^{\frac{1}{2}}$,

$$\begin{aligned}
\big(\beta\mathbf{I} + (\mathbf{I} - \boldsymbol{P})\big)^{-1} &= (\beta\mathbf{I} + \mathbf{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\mathbf{\Pi}^{\frac{1}{2}})^{-1} \\
&= (\beta\mathbf{I})^{-1} - (\beta\mathbf{I})^{-1}\mathbf{\Pi}^{-\frac{1}{2}}\big(\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}}\mathbf{\Pi}^{\frac{1}{2}}(\beta\mathbf{I})^{-1}\mathbf{\Pi}^{-\frac{1}{2}}\big)^{-1}\tilde{\boldsymbol{\mathcal{L}}}\mathbf{\Pi}^{\frac{1}{2}}(\beta\mathbf{I})^{-1} \\
&= (\beta\mathbf{I})^{-1} - (\beta\mathbf{I})^{-1}\mathbf{\Pi}^{-\frac{1}{2}}\big(\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I})^{-1}\big)^{-1}\tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I})^{-1}\mathbf{\Pi}^{\frac{1}{2}} \\
&= (\beta\mathbf{I})^{-1} - (\beta\mathbf{I})^{-1}\mathbf{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I})^{-1}\big(\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I})^{-1}\big)^{-1}\mathbf{\Pi}^{\frac{1}{2}} \\
&= \frac{1}{\beta}\big(\mathbf{I} - \mathbf{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1}\mathbf{\Pi}^{\frac{1}{2}}\big).
\end{aligned}$$

On the second line, we used the generalized identity for the inverse of a sum of matrices $(\mathbf{A} + \mathbf{UBV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{I} + \mathbf{BVA}^{-1}\mathbf{U})^{-1}\mathbf{BVA}^{-1}$ where $\mathbf{A}$ is nonsingular and $\mathbf{U}$, $\mathbf{B}$, and $\mathbf{V}$ are square matrices in our case. On the fourth line, we used the identity $(\mathbf{I}+\mathbf{C})^{-1}\mathbf{C} = \mathbf{C}(\mathbf{I}+\mathbf{C})^{-1}$ where for any $\mathbf{C}$, $(\mathbf{I}+\mathbf{C})^{-1}$ is nonsingular [30].

The personalized PageRank for coevolutionary digraphs is calculated as

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) = \mathbf{s}\big(\mathbf{I} - \mathbf{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1}\mathbf{\Pi}^{\frac{1}{2}}\big).$$

since $(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1}$ is nonsingular (by Theorem 1.2.17 in page 54, [26]). What happens when $\beta \to 0$? We know the linear system should reduce to the case whereby the solution is the stationary vector $\boldsymbol{\pi}$ if we assume that the coevolutionary digraph is irreducible.

Since $\tilde{\boldsymbol{\mathcal{Z}}} = \tilde{\boldsymbol{\mathcal{L}}}^+$ is the Moore-Penrose pseudoinverse of $\tilde{\boldsymbol{\mathcal{L}}}$,

$$\begin{aligned}
\tilde{\boldsymbol{\mathcal{L}}} &= \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\tilde{\boldsymbol{\mathcal{L}}} \\
\tilde{\boldsymbol{\mathcal{L}}} + \boldsymbol{\mathcal{E}} &= \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}}) \\
\boldsymbol{\mathcal{E}} &= \beta\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}.
\end{aligned}$$

Then,

$$\begin{aligned}
\tilde{\boldsymbol{\mathcal{L}}} + \boldsymbol{\mathcal{E}} &= \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}}) \\
(\tilde{\boldsymbol{\mathcal{L}}} + \boldsymbol{\mathcal{E}})(\beta\mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1} &= \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}
\end{aligned}$$

$$\begin{aligned}\tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1} &= \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}} - \boldsymbol{\mathcal{E}}(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1}\\ &= \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\big(\mathbf{I}-\beta(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1}\big).\end{aligned}$$

We rewrite the personalized PageRank for coevolutionary digraphs as follows

$$\begin{aligned}\boldsymbol{\psi}_\alpha(\mathbf{s}) &= \mathbf{s}\big(\mathbf{I}-\boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\big(\mathbf{I}-\beta(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1}\big)\boldsymbol{\Pi}^{\frac{1}{2}}\big)\\
&= \mathbf{s}\big(\mathbf{I}-\boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\boldsymbol{\Pi}^{\frac{1}{2}}+\beta\boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\boldsymbol{\Pi}^{\frac{1}{2}}\boldsymbol{\Pi}^{-\frac{1}{2}}(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1}\boldsymbol{\Pi}^{\frac{1}{2}}\big)\\
&= \mathbf{s}\big(\mathbf{I}-\boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\boldsymbol{\Pi}^{\frac{1}{2}}+\beta\boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\boldsymbol{\Pi}^{\frac{1}{2}}\big(\beta\boldsymbol{\Pi}^{-\frac{1}{2}}\boldsymbol{\Pi}^{\frac{1}{2}}+\boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\boldsymbol{\Pi}^{\frac{1}{2}}\big)^{-1}\big)\\
&= \mathbf{s}\big(\mathbf{I}-\boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\boldsymbol{\Pi}^{\frac{1}{2}}\big)+\beta\mathbf{s}\big(\boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\boldsymbol{\Pi}^{\frac{1}{2}}\big(\beta\mathbf{I}+\mathbf{I}-\boldsymbol{P}\big)^{-1}\big)\\
&= \mathbf{s}\big(\mathbf{I}-\boldsymbol{\Pi}^{-\frac{1}{2}}(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\boldsymbol{\Pi}^{\frac{1}{2}}\big)+\beta\mathbf{s}\big(\boldsymbol{\Pi}^{-\frac{1}{2}}(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\boldsymbol{\Pi}^{\frac{1}{2}}\big(\beta\mathbf{I}+\mathbf{I}-\boldsymbol{P}\big)^{-1}\big)\\
&= \boldsymbol{\pi}+\beta(\mathbf{s}-\boldsymbol{\pi})\big(\beta\mathbf{I}+\mathbf{I}-\boldsymbol{P}\big)^{-1}\end{aligned}$$

where $\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}} = (\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})$. We need to show that $\mathbf{s}\big(\mathbf{I}-\boldsymbol{\Pi}^{-\frac{1}{2}}(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\boldsymbol{\Pi}^{\frac{1}{2}}\big)=\boldsymbol{\pi}$ and $\mathbf{s}\big(\boldsymbol{\Pi}^{-\frac{1}{2}}(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\boldsymbol{\Pi}^{\frac{1}{2}}\big)=\mathbf{s}-\boldsymbol{\pi}$.

First, we obtain by direct calculation that

$$\mathbf{I}-\boldsymbol{\Pi}^{-\frac{1}{2}}(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\boldsymbol{\Pi}^{\frac{1}{2}} = \begin{pmatrix}\pi_1 & \pi_2 & \pi_3 & \cdots & \pi_n\\ \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_n\\ \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_n\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_n\end{pmatrix}$$

where $\tilde{\boldsymbol{\mathcal{J}}} = (\boldsymbol{\pi}^{1/2})^T\boldsymbol{\pi}^{1/2}$. Since $\mathbf{s}=(s_i)_{i=1}^n$ with $\sum_i s_i = 1$, we can calculate the following $\mathbf{s}\big(\mathbf{I}-\boldsymbol{\Pi}^{-\frac{1}{2}}(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\boldsymbol{\Pi}^{\frac{1}{2}}\big) = (\sum_i \pi_j s_i)_{j=1}^n = (\pi_j\sum_i s_i)_{j=1}^n = (\pi_j)_{j=1}^n = \boldsymbol{\pi}$.

Second, let $\mathbf{s}=(s_1,s_2,s_3,\ldots,s_n)$ and $\boldsymbol{\pi}=(\pi_1,\pi_2,\pi_3,\ldots,\pi_n)$. Direct calculation shows that $\mathbf{s}\big(\boldsymbol{\Pi}^{-\frac{1}{2}}(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\boldsymbol{\Pi}^{\frac{1}{2}}\big) = (s_1-\pi_1, s_2-\pi_2, s_3-\pi_3,\ldots,s_n-\pi_n) = \mathbf{s}-\boldsymbol{\pi}$. Note that the product of the row vector $\mathbf{s}$ with the first column of the matrix $\boldsymbol{\Pi}^{-\frac{1}{2}}(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\boldsymbol{\Pi}^{\frac{1}{2}}$ is

$$\begin{aligned}&(s_1,s_2,s_3,\ldots,s_n)(1-\pi_1,-\pi_1,-\pi_1,\ldots,-\pi_1)^T\\ &= s_1(1-\pi_1)+s_2(-\pi_1)+s_3(-\pi_1)+\cdots+s_n(-\pi_1)\\ &= s_1-\pi_1\end{aligned}$$

whereby the product of $\mathbf{s}$ with the ith-column of $\mathbf{\Pi}^{-\frac{1}{2}}(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\mathbf{\Pi}^{\frac{1}{2}}$ is $s_i - \pi_i$.

We have calculated that

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) = \boldsymbol{\pi} + \beta(\mathbf{s}-\boldsymbol{\pi})\big(\beta\mathbf{I}+\mathbf{I}-\boldsymbol{P}\big)^{-1}. \tag{A.5}$$

Since $\beta\mathbf{I}+\mathbf{I}-\boldsymbol{P}$ is nonsingular, then

$$\boldsymbol{\psi}_\alpha(\mathbf{s})\big(\beta\mathbf{I}+\mathbf{I}-\boldsymbol{P}\big) = \boldsymbol{\pi}(\beta\mathbf{I}+\mathbf{I}-\boldsymbol{P}) + \beta(\mathbf{s}-\boldsymbol{\pi}).$$

Note that when we let $\alpha \to 0 \Rightarrow \beta \to 0$, we obtain

$$\begin{aligned}\boldsymbol{\psi}_0(\mathbf{s})\big(\mathbf{I}-\boldsymbol{P}\big) &= \boldsymbol{\pi}(\mathbf{I}-\boldsymbol{P}) + \tilde{\mathbf{0}}\\ &= \boldsymbol{\pi}(\mathbf{I}-\boldsymbol{P})\end{aligned}$$

where $\tilde{\mathbf{0}}$ is the zero row vector. As such, $\boldsymbol{\psi}_0(\mathbf{s}) = \boldsymbol{\pi}$ for any $\mathbf{s}$.

Can we calculate $\boldsymbol{\psi}_\alpha(\mathbf{s}) = \mathbf{s}\big(\mathbf{I}-\mathbf{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\big(\mathbf{I}-\beta(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1}\big)\mathbf{\Pi}^{\frac{1}{2}}\big)$ directly? We need to be able to calculate $\boldsymbol{\psi}_0(\mathbf{s})$ also. Let $\mathbf{N}_\beta = \beta(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1}$. This means calculating $\mathbf{N}_0$ also. $\mathbf{N}_\beta$ is nonsingular. Then,

$$\begin{aligned}\mathbf{N}_\beta &= \beta(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1}\\ \beta\mathbf{I}\mathbf{N}_\beta + \mathbf{N}_\beta\tilde{\boldsymbol{\mathcal{L}}} &= \beta\mathbf{I}.\end{aligned}$$

Assume $\mathbf{N}_0$ is nonsingular. As $\beta \to 0$, we have

$$\begin{aligned}\mathbf{0}\mathbf{N}_0 + \mathbf{N}_0\tilde{\boldsymbol{\mathcal{L}}} &= \mathbf{0}\\ \mathbf{N}_0\tilde{\boldsymbol{\mathcal{L}}} &= \mathbf{0}.\end{aligned}$$

We know $\tilde{\boldsymbol{\mathcal{L}}}$ is singular. But for the last equality $\mathbf{N}_0\tilde{\boldsymbol{\mathcal{L}}} = \mathbf{0}$, the original statement that $\mathbf{N}_0$ is nonsingular would be a contradiction. $\mathbf{N}_0$ must be singular.

We can apply the theoretical result involving several identities relating to the normalized digraph Laplacians [29] and obtain the solution for $\mathbf{N}_0\tilde{\boldsymbol{\mathcal{L}}} = \mathbf{0}$ as $\mathbf{N}_0 = \tilde{\boldsymbol{\mathcal{J}}}^k = \tilde{\boldsymbol{\mathcal{J}}}$, $k = 1, 2, 3, \ldots$ (by repeatedly applying the identity $\tilde{\boldsymbol{\mathcal{J}}}^2 = \tilde{\boldsymbol{\mathcal{J}}}$) or $\mathbf{N}_0 = \mathbf{0}$ (the zero matrix). Here, $\tilde{\boldsymbol{\mathcal{J}}}\tilde{\boldsymbol{\mathcal{L}}} = \mathbf{0}$ (by Lemma 1 in [29]) and $\mathbf{0}\tilde{\boldsymbol{\mathcal{L}}} = \mathbf{0}$. We furnish a direct calculation for $\tilde{\boldsymbol{\mathcal{J}}}\tilde{\boldsymbol{\mathcal{L}}}$ here since it is sketched only in [29]. Let $\mathbf{J}$ be the square matrix of ones and $\mathbf{e}$ the row vector of ones. Then,

$$\tilde{\boldsymbol{\mathcal{J}}}\tilde{\boldsymbol{\mathcal{L}}} = (\mathbf{\Pi}^{\frac{1}{2}}\mathbf{J}\mathbf{\Pi}^{\frac{1}{2}})(\mathbf{\Pi}^{\frac{1}{2}}(\mathbf{I}-\boldsymbol{P})\mathbf{\Pi}^{-\frac{1}{2}})$$

$$
\begin{aligned}
&= \mathbf{\Pi}^{\frac{1}{2}}\mathbf{J}\mathbf{\Pi}(\mathbf{I}-\boldsymbol{P})\mathbf{\Pi}^{-\frac{1}{2}} \\
&= \mathbf{\Pi}^{\frac{1}{2}}\mathbf{e}^T\boldsymbol{\pi}(\mathbf{I}-\boldsymbol{P})\mathbf{\Pi}^{-\frac{1}{2}} \\
&= \mathbf{0}
\end{aligned}
$$

since $\boldsymbol{\pi}(\mathbf{I}-\boldsymbol{P}) = \boldsymbol{\pi} - \boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi} - \boldsymbol{\pi} = \tilde{\mathbf{0}}$ and $\mathbf{e}^T\tilde{\mathbf{0}} = \mathbf{0}$ where $\tilde{\mathbf{0}}$ is the zero row vector and $\mathbf{0}$ is the zero matrix.

However, we can establish what $\mathbf{N}_0$ is, directly. First we rewrite the expression $(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1}$ as follows

$$
\begin{aligned}
(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1} &= \big(\beta\mathbf{I}+(\tilde{\boldsymbol{\mathcal{L}}}+\tilde{\boldsymbol{\mathcal{J}}})-\tilde{\boldsymbol{\mathcal{J}}}\big)^{-1} \\
&= \Big((\tilde{\boldsymbol{\mathcal{L}}}+\tilde{\boldsymbol{\mathcal{J}}})\big(\mathbf{I}+(\tilde{\boldsymbol{\mathcal{L}}}+\tilde{\boldsymbol{\mathcal{J}}})^{-1}(\beta\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\big)\Big)^{-1} \\
&= \big(\tilde{\boldsymbol{\mathcal{L}}}+\tilde{\boldsymbol{\mathcal{J}}}\big)^{-1}\big(\mathbf{I}+(\tilde{\boldsymbol{\mathcal{L}}}+\tilde{\boldsymbol{\mathcal{J}}})^{-1}(\beta\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\big)^{-1} \\
&= \big(\tilde{\boldsymbol{\mathcal{Z}}}+\tilde{\boldsymbol{\mathcal{J}}}\big)\big(\mathbf{I}+(\tilde{\boldsymbol{\mathcal{Z}}}+\tilde{\boldsymbol{\mathcal{J}}})(\beta\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\big)^{-1} \\
&= \big(\tilde{\boldsymbol{\mathcal{Z}}}+\tilde{\boldsymbol{\mathcal{J}}}\big)\big(\mathbf{I}+\beta(\tilde{\boldsymbol{\mathcal{Z}}}+\tilde{\boldsymbol{\mathcal{J}}})-\tilde{\boldsymbol{\mathcal{Z}}}\tilde{\boldsymbol{\mathcal{J}}}-\tilde{\boldsymbol{\mathcal{J}}}\tilde{\boldsymbol{\mathcal{J}}}\big)^{-1} \\
&= \big(\tilde{\boldsymbol{\mathcal{Z}}}+\tilde{\boldsymbol{\mathcal{J}}}\big)\big(\beta(\tilde{\boldsymbol{\mathcal{Z}}}+\tilde{\boldsymbol{\mathcal{J}}})+(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\big)^{-1},
\end{aligned}
$$

using identities $(\tilde{\boldsymbol{\mathcal{Z}}}+\tilde{\boldsymbol{\mathcal{J}}}) = (\tilde{\boldsymbol{\mathcal{L}}}+\tilde{\boldsymbol{\mathcal{J}}})^{-1}$ on the fourth line (by Theorem 1 in [29]), $\tilde{\boldsymbol{\mathcal{Z}}}\tilde{\boldsymbol{\mathcal{J}}} = \mathbf{0}$ and $\tilde{\boldsymbol{\mathcal{J}}}\tilde{\boldsymbol{\mathcal{J}}} = \tilde{\boldsymbol{\mathcal{J}}}$ on the last line (by Lemma 1 in [29]).

Since $\mathbf{N}_\beta = \beta(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1}$ is nonsingular, then

$$
\begin{aligned}
\mathbf{N}_\beta &= \beta\big(\tilde{\boldsymbol{\mathcal{Z}}}+\tilde{\boldsymbol{\mathcal{J}}}\big)\big(\beta(\tilde{\boldsymbol{\mathcal{Z}}}+\tilde{\boldsymbol{\mathcal{J}}})+(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\big)^{-1} \\
\mathbf{N}_\beta\big(\beta(\tilde{\boldsymbol{\mathcal{Z}}}+\tilde{\boldsymbol{\mathcal{J}}})+(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\big) &= \beta\big(\tilde{\boldsymbol{\mathcal{Z}}}+\tilde{\boldsymbol{\mathcal{J}}}\big).
\end{aligned}
$$

As $\beta \to 0$, we have

$$
\begin{aligned}
\mathbf{N}_0\big(\mathbf{0}+(\mathbf{I}-\tilde{\boldsymbol{\mathcal{J}}})\big) &= \mathbf{0} \\
\mathbf{N}_0 &= \mathbf{N}_0\tilde{\boldsymbol{\mathcal{J}}}.
\end{aligned}
$$

Given that $\tilde{\boldsymbol{\mathcal{J}}}$ is singular, the solutions to the equality $\mathbf{N}_0 = \mathbf{N}_0\tilde{\boldsymbol{\mathcal{J}}}$ are $\mathbf{N}_0 = \mathbf{0}$ and $\mathbf{N}_0 = \tilde{\boldsymbol{\mathcal{J}}}^k = \tilde{\boldsymbol{\mathcal{J}}}$, $k = 1, 2, 3, \ldots$.

As $\alpha \to 0 \Rightarrow \beta \to 0$, we can then reduce the equation $\tilde{\boldsymbol{\mathcal{L}}}(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1} = \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\big(\mathbf{I}-\beta(\beta\mathbf{I}+\tilde{\boldsymbol{\mathcal{L}}})^{-1}\big)$ as follows

$$
\tilde{\boldsymbol{\mathcal{L}}}(\tilde{\boldsymbol{\mathcal{L}}})^{+} = \tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}(\mathbf{I}-\mathbf{N}_0)
$$

$$= \tilde{\mathcal{L}}\tilde{\mathcal{Z}}(\mathbf{I} - \tilde{\mathcal{J}})$$
$$= \tilde{\mathcal{L}}\tilde{\mathcal{Z}},$$

given that $\tilde{\mathcal{L}}\tilde{\mathcal{Z}} = \mathbf{I} - \tilde{\mathcal{J}}$ and $\tilde{\mathcal{L}}\tilde{\mathcal{Z}}\tilde{\mathcal{L}} = \tilde{\mathcal{L}}$. The conclusion is the same if we consider $\mathbf{N}_0 = \mathbf{0}$. This implies that

$$\lim_{\beta \to 0} \tilde{\mathcal{L}}(\beta \mathbf{I} + \tilde{\mathcal{L}})^{-1} = \tilde{\mathcal{L}}(\tilde{\mathcal{L}})^{+}$$
$$= \tilde{\mathcal{L}}\tilde{\mathcal{Z}}.$$

For small values of α (subsequently small values of β), we can use this approximation

$$\tilde{\mathcal{L}}(\beta \mathbf{I} + \tilde{\mathcal{L}})^{-1} \approx \tilde{\mathcal{L}}\tilde{\mathcal{Z}} = \mathbf{I} - \tilde{\mathcal{J}},$$

with equality when $\beta = 0$. We can approximate the personalized PageRank for irreducible coevolutionary digraphs as

$$\boldsymbol{\psi}_\alpha(\mathbf{s}) \approx \mathbf{s}\big(\mathbf{I} - \mathbf{\Pi}^{-\frac{1}{2}}(\mathbf{I} - \tilde{\mathcal{J}})\mathbf{\Pi}^{\frac{1}{2}}\big) = \boldsymbol{\pi}.$$

with equality, i.e., $\boldsymbol{\psi}_0(\mathbf{s}) = \boldsymbol{\pi}$, for any $\mathbf{s}$. □

Lemma 4. *Let a CMC operating on irreducible $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$ be associated with a probability transition matrix $\boldsymbol{P}$, stationary distribution $\boldsymbol{\pi}$, and the fundamental matrix Z. Correspondingly, the personalized PageRank on $D_C(V_{S(n)})$ is a CMC with probability transition matrix $\boldsymbol{P}_{PR}$ and stationary distribution $\boldsymbol{\psi}_\alpha(\mathbf{s})$. Then*

$$\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}) = -\boldsymbol{\psi}_\alpha(\mathbf{s})\big(\mathbf{I} - \boldsymbol{P}\big)Z$$
$$= \beta(\boldsymbol{\pi} - \mathbf{s})\big(\beta\mathbf{I} + (\mathbf{I} - \boldsymbol{P})\big)^{-1}.$$

Proof: The following makes use of results of Perturbation theory applied to finite Markov chains in [31]. Let $\boldsymbol{\pi}_A = \boldsymbol{\pi}$, $\boldsymbol{P}_A = \boldsymbol{P}$, $\boldsymbol{\pi}_B = \boldsymbol{\psi}_\alpha(\mathbf{s})$, and $\boldsymbol{P}_B = \boldsymbol{P}_{\mathrm{PR}}$. The distance between $\boldsymbol{P}$ and $\boldsymbol{P}_{\mathrm{PR}}$ is defined as

$$\mathbf{U}_{AB} = (\boldsymbol{P}_{\mathrm{PR}} - \boldsymbol{P})Z$$

with the fundamental matrix $Z = Z_A$. Applying (Theorems 1 and 2 in [31]), we obtain

$$\begin{aligned}
\boldsymbol{\psi}_\alpha(\mathbf{s}) &= \boldsymbol{\pi}(\mathbf{I} - \mathbf{U}_{AB})^{-1} \\
\boldsymbol{\psi}_\alpha(\mathbf{s}) - \boldsymbol{\psi}_\alpha(\mathbf{s})\mathbf{U}_{AB} &= \boldsymbol{\pi} \\
\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}) &= -\boldsymbol{\psi}_\alpha(\mathbf{s})\mathbf{U}_{AB} \\
&= \boldsymbol{\psi}_\alpha(\mathbf{s})(\boldsymbol{P} - \boldsymbol{P}_{\mathrm{PR}})Z \\
&= -\boldsymbol{\psi}_\alpha(\mathbf{s})(\mathbf{I} - \boldsymbol{P})Z
\end{aligned}$$

making use of $\boldsymbol{\psi}_\alpha(\mathbf{s})\boldsymbol{P}_{\mathrm{PR}} = \boldsymbol{\psi}_\alpha(\mathbf{s})$.

We show that the vector $\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})$ is invariant under multiplication of the matrix $(\mathbf{I} - \boldsymbol{P})Z$, i.e., $(\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}))(\mathbf{I} - \boldsymbol{P})Z = \boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})$. Furthermore, $\big((\mathbf{I} - \boldsymbol{P})Z\big)^k = (\mathbf{I} - \boldsymbol{P})Z$, $k = 1, 2, 3, \ldots$. First,

$$\begin{aligned}
(\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}))(\mathbf{I} - \boldsymbol{P})Z &= \boldsymbol{\pi}(\mathbf{I} - \boldsymbol{P})Z - \boldsymbol{\psi}_\alpha(\mathbf{s})(\mathbf{I} - \boldsymbol{P})Z \\
&= \tilde{\mathbf{0}} - \boldsymbol{\psi}_\alpha(\mathbf{s})(\mathbf{I} - \boldsymbol{P})Z \\
&= \boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}).
\end{aligned}$$

where $\tilde{\mathbf{0}}$ is the zero row vector. Second, using identities $\tilde{\boldsymbol{\mathcal{L}}} = \boldsymbol{\Pi}^{\frac{1}{2}}(\mathbf{I} - \boldsymbol{P})\boldsymbol{\Pi}^{-\frac{1}{2}}$ and $\tilde{\boldsymbol{\mathcal{Z}}} = \boldsymbol{\Pi}^{\frac{1}{2}} Z \boldsymbol{\Pi}^{-\frac{1}{2}}$ [29], we obtain $(\mathbf{I} - \boldsymbol{P})Z = \boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\boldsymbol{\Pi}^{\frac{1}{2}}$. Then,

$$\begin{aligned}
\big((\mathbf{I} - \boldsymbol{P})Z\big)\big((\mathbf{I} - \boldsymbol{P})Z\big) &= (\boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\boldsymbol{\Pi}^{\frac{1}{2}})(\boldsymbol{\Pi}^{-\frac{1}{2}}\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\boldsymbol{\Pi}^{\frac{1}{2}}) \\
&= (\mathbf{I} - \boldsymbol{P})Z
\end{aligned}$$

making use of $\tilde{\boldsymbol{\mathcal{L}}}\tilde{\boldsymbol{\mathcal{Z}}}\tilde{\boldsymbol{\mathcal{L}}} = \tilde{\boldsymbol{\mathcal{L}}}$. Applying this equality repeatedly we obtain $\big((\mathbf{I} - \boldsymbol{P})Z\big)^k = (\mathbf{I} - \boldsymbol{P})Z$, $k = 1, 2, 3, \ldots$.

Applying Equation A.5 from the proof section of Theorem 3, we obtain

$$\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}) = \beta(\boldsymbol{\pi} - \mathbf{s})\big(\beta\mathbf{I} + \mathbf{I} - \boldsymbol{P}\big)^{-1}.$$

Multiplying $(\mathbf{I} - \boldsymbol{P})Z$ on both sides of the equality

$$\begin{aligned}
(\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}))(\mathbf{I} - \boldsymbol{P})Z &= \beta(\boldsymbol{\pi} - \mathbf{s})\big(\beta\mathbf{I} + \mathbf{I} - \boldsymbol{P}\big)^{-1}(\mathbf{I} - \boldsymbol{P})Z \\
\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}) &= \beta\boldsymbol{\pi}(\mathbf{I} - \boldsymbol{P})Z - \beta\mathbf{s}(\beta\mathbf{I} + \mathbf{I} - \boldsymbol{P})^{-1}(\mathbf{I} - \boldsymbol{P})Z
\end{aligned}$$

$$= \tilde{\mathbf{0}} - \beta \mathbf{s}(\beta \mathbf{I} + \mathbf{I} - \boldsymbol{P})^{-1}(\mathbf{I} - \boldsymbol{P})Z$$
$$= -\boldsymbol{\psi}_\alpha(\mathbf{s})(\mathbf{I} - \boldsymbol{P})Z,$$

making use of Equation A.2 from Lemma 1, which completes the proof. □

Corollary 5. *For $\boldsymbol{\pi}$ and $\boldsymbol{\psi}_\alpha(\mathbf{s})$ associated with CMC on an irreducible $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$, the following inequality is given for restart probabilities $\alpha_1 \leq \alpha_2$*

$$||\boldsymbol{\pi} - \boldsymbol{\psi}_{\alpha_1}(\mathbf{s})|| \leq ||\boldsymbol{\pi} - \boldsymbol{\psi}_{\alpha_2}(\mathbf{s})|| \tag{A.6}$$

where $\alpha_1, \alpha_2 \in (0, 1)$.

Proof: First, we rewrite the equation as follows

$$\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}) = \beta(\boldsymbol{\pi} - \mathbf{s})(\beta \mathbf{I} + \mathbf{I} - \boldsymbol{P})^{-1}$$
$$\mathbf{y} = (\boldsymbol{\pi} - \mathbf{s})\mathbf{A}^{-1}$$

where $\mathbf{y} = \boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})$ and $\mathbf{A} = \mathbf{I} + \frac{1}{\beta}(\mathbf{I} - \boldsymbol{P})$. $\mathbf{A}^{-1}$ is by definition a nonsingular M-matrix and as such is inverse-positive $\mathbf{A}^{-1} \geq \mathbf{0}$ (Theorem 2.3, Chapter 6, [27]) with positive $|| \star ||$. Furthermore, since each row sum of $\mathbf{A}$ is equal to one, then $\mathbf{A}$ is a uniformly strictly diagonally dominant M-matrix and that $\mathbf{A}^{-1}$ is a row stochastic matrix (Corollary 2 in [43]). We have three cases when: (a) $\alpha = 0$, (b) $\alpha = 1$, and (c) $\alpha \in (0, 1)$. We evaluate these cases individually using several results in the proof of Theorem 3.

For (a), Theorem 3 implies that as $\alpha \to 0 \Rightarrow \beta \to 0$, we have $\lim_{\beta \to 0} ||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})|| = 0$. We can show this directly here. We rewrite $\mathbf{A}^{-1}$ as

$$\mathbf{A}^{-1} = \big(\mathbf{I} + \frac{1}{\beta}(\mathbf{I} - \boldsymbol{P})\big)^{-1}$$
$$= \mathbf{I} - \boldsymbol{\Pi}^{-\frac{1}{2}} \tilde{\boldsymbol{\mathcal{L}}}(\beta \mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1} \boldsymbol{\Pi}^{\frac{1}{2}}.$$

since $\big(\beta \mathbf{I} + (\mathbf{I} - \boldsymbol{P})\big)^{-1} = \frac{1}{\beta}\big(\mathbf{I} - \boldsymbol{\Pi}^{-\frac{1}{2}} \tilde{\boldsymbol{\mathcal{L}}}(\beta \mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1} \boldsymbol{\Pi}^{\frac{1}{2}}\big)$. Let $\mathbf{y}(0) = \lim_{\beta \to 0}(\boldsymbol{\pi} - \mathbf{s})\mathbf{A}^{-1}$. As $\alpha \to 0 \Rightarrow \beta \to 0$, we obtain

$$\mathbf{y}(0) = \lim_{\beta \to 0} (\boldsymbol{\pi} - \mathbf{s})\mathbf{A}^{-1}$$

$$
\begin{aligned}
&= (\boldsymbol{\pi} - \mathbf{s}) \lim_{\beta\to 0} (\mathbf{I} - \mathbf{\Pi}^{-\frac{1}{2}} \tilde{\boldsymbol{\mathcal{L}}}(\beta \mathbf{I} + \tilde{\boldsymbol{\mathcal{L}}})^{-1} \mathbf{\Pi}^{\frac{1}{2}}) \\
&= (\boldsymbol{\pi} - \mathbf{s}) \mathbf{e}^T \boldsymbol{\pi} \\
&= \tilde{\mathbf{0}}
\end{aligned}
$$

making use of $\mathbf{I} - \mathbf{\Pi}^{-\frac{1}{2}} \tilde{\boldsymbol{\mathcal{L}}} \tilde{\boldsymbol{\mathcal{Z}}} \mathbf{\Pi}^{\frac{1}{2}} = \mathbf{I} - \mathbf{\Pi}^{-\frac{1}{2}} (\mathbf{I} - \tilde{\boldsymbol{\mathcal{J}}}) \mathbf{\Pi}^{\frac{1}{2}} = \mathbf{e}^T \boldsymbol{\pi}$. So, $||\mathbf{y}(0)|| = 0$.

For (b), as $\alpha \to 1 \Rightarrow \beta \to \infty$, we have $\lim_{\beta\to\infty} ||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})|| = ||\boldsymbol{\pi} - \mathbf{s}||$ since $\boldsymbol{\psi}_1(\mathbf{s}) = \mathbf{s}$, given that $\lim_{\beta\to\infty} \mathbf{A}^{-1} = \mathbf{I}^{-1} = \mathbf{I}$ and $\boldsymbol{\psi}_\alpha(\mathbf{s}) = \mathbf{s}\mathbf{A}^{-1}$. We can show this directly as follows. Let $\mathbf{y}(1) = \lim_{\beta\to\infty} (\boldsymbol{\pi} - \mathbf{s}) \mathbf{A}^{-1}$. Then,

$$
\begin{aligned}
\mathbf{y}(1) &= \lim_{\beta\to\infty} (\boldsymbol{\pi} - \mathbf{s}) \mathbf{A}^{-1} \\
&= (\boldsymbol{\pi} - \mathbf{s}) \lim_{\beta\to\infty} \mathbf{A}^{-1} \\
&= \boldsymbol{\pi} - \mathbf{s}.
\end{aligned}
$$

So, $||\mathbf{y}(1)|| = ||\boldsymbol{\pi} - \mathbf{s}||$.

For (c), note that $\mathbf{y}$ and $\mathbf{A}$ consist of entries that are differentiable functions of a real variable β [22] whereby

$$
\begin{aligned}
\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\beta} &= (\boldsymbol{\pi} - \mathbf{s}) \frac{\mathrm{d}}{\mathrm{d}\beta} \mathbf{A}^{-1} \\
&= \frac{1}{\beta^2} (\boldsymbol{\pi} - \mathbf{s}) \mathbf{A}^{-1} (\mathbf{I} - \boldsymbol{P}) \mathbf{A}^{-1}
\end{aligned}
$$

since

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\beta} \mathbf{A} &= \frac{\mathrm{d}}{\mathrm{d}\beta} \big(\mathbf{I} + \frac{1}{\beta} (\mathbf{I} - \boldsymbol{P}) \big) \\
&= -\frac{1}{\beta^2} (\mathbf{I} - \boldsymbol{P})
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\beta} \mathbf{A}^{-1} &= -\mathbf{A}^{-1} (\frac{\mathrm{d}}{\mathrm{d}\beta} \mathbf{A}) \mathbf{A}^{-1} \\
&= \frac{1}{\beta^2} \mathbf{A}^{-1} (\mathbf{I} - \boldsymbol{P}) \mathbf{A}^{-1}.
\end{aligned}
$$

We note that

$$
(\boldsymbol{\pi} - \mathbf{s}) \mathbf{A}^{-1} (\mathbf{I} - \boldsymbol{P}) \mathbf{A}^{-1} = \big(\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}) \big) (\mathbf{I} - \boldsymbol{P}) \mathbf{A}^{-1}
$$

$$= (\boldsymbol{\psi}_\alpha(\mathbf{s})\boldsymbol{P} - \boldsymbol{\psi}_\alpha(\mathbf{s}))\mathbf{A}^{-1}.$$

Given that $\lim_{\beta\to\infty} \mathbf{A}^{-1} = \mathbf{I}$, we obtain

$$\lim_{\beta\to\infty} (\boldsymbol{\pi} - \mathbf{s})\mathbf{A}^{-1}(\mathbf{I} - \boldsymbol{P})\mathbf{A}^{-1} = (\boldsymbol{\pi} - \mathbf{s})(\mathbf{I} - \boldsymbol{P})$$
$$= \mathbf{s}\boldsymbol{P} - \mathbf{s}.$$

As $\alpha \to 1 \Rightarrow \beta \to \infty$, $\lim_{\beta\to\infty} \left|\left|\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\beta}\right|\right| = \lim_{\beta\to\infty} \frac{1}{\beta^2}||\mathbf{s}\boldsymbol{P} - \mathbf{s}|| = 0.$

$\left|\left|\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\beta}\right|\right|$ is positive for $\alpha \in (0, 1)$ and goes to zero as $\alpha \to 1$. Furthermore, we know that $\mathbf{y}(\alpha)$ is bounded entrywise with $||\mathbf{y}(0)|| = 0$ and $||\mathbf{y}(1)|| = ||\boldsymbol{\pi} - \mathbf{s}||$. Since $||\mathbf{y}(\alpha)||$ is monotonically increasing in the interval of α for $(0, 1)$, it follows that $||\boldsymbol{\pi} - \boldsymbol{\psi}_{\alpha_1}(\mathbf{s})| \leq ||\boldsymbol{\pi} - \boldsymbol{\psi}_{\alpha_2}(\mathbf{s})||$ for $\alpha_1 \leq \alpha_2$ where $\alpha_1, \alpha_2 \in (0, 1)$. □

Corollary 6. *Associated to each irreducible* $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$ *is CMC with stationary distribution* $\boldsymbol{\psi}_\alpha(\mathbf{s})$ *for* $\alpha \in (0, 1)$ *and* $\boldsymbol{\pi}$ *with the following perturbation bound*

$$||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty \leq ||\boldsymbol{\pi} - \mathbf{s}||_\infty \tag{A.7}$$

with equality when $\alpha = 1$.

Proof: Since $||(\beta\mathbf{I} + \mathbf{I} - \boldsymbol{P})^{-1}||_\infty = \frac{1}{\beta}$ from Lemma 2 and from the property of induced matrix norm [22], we obtain

$$||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_\infty = ||\beta(\boldsymbol{\pi} - \mathbf{s})(\beta\mathbf{I} + \mathbf{I} - \boldsymbol{P})^{-1}||_\infty$$
$$\leq \beta||\boldsymbol{\pi} - \mathbf{s}||_\infty \, ||(\beta\mathbf{I} + \mathbf{I} - \boldsymbol{P})^{-1}||_\infty$$
$$= ||\boldsymbol{\pi} - \mathbf{s}||_\infty.$$

Equality is obtained as $\alpha = 1$ since $\boldsymbol{\psi}_1(\mathbf{s}) = \mathbf{s}$. □

Note: Lemma 4 expresses the difference between stationary and PageRank vectors of a CMC explicitly, which is followed by Corollary 5 that shows for suitable vector norms, $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||$ is directly proportional to the restart probability α. This can be seen from the expression $\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s}) = (\boldsymbol{\pi} - \mathbf{s})\mathbf{A}^{-1}$ where $\mathbf{A}^{-1} = \left(\mathbf{I} + \frac{1}{\beta}(\mathbf{I} - \boldsymbol{P})\right)^{-1}$. $\mathbf{A}^{-1}$ is a row stochastic matrix that changes from

$\mathbf{e}^T\boldsymbol{\pi}$ (as $\alpha \to 0$) to $\mathbf{I}$ (as $\alpha \to 1$). Although it is not possible to find a general expression of $\mathbf{A}^{-1}$ for $\alpha \in (0,1)$, for $\mathbf{A}$ being a nonsingular M-matrix that is diagonally dominant, $\mathbf{A}^{-1}$ is diagonally dominant of its column entries (i.e, for $\mathbf{B} = \mathbf{A}^{-1}$, $|b_{ii}| > |b_{ij}|,\ j \neq i,\ i = 1, 2, 3, \ldots, n$) [44]. Furthermore, $\boldsymbol{\pi}\mathbf{A}^{-1} = \boldsymbol{\pi}$ from Equation A.5. One can observe that $||(\boldsymbol{\pi} - \mathbf{s})\mathbf{A}^{-1}|| = ||\boldsymbol{\pi} - \mathbf{s}\mathbf{A}^{-1}||$ monotonically increases as $\alpha \to 1$. For example, $||(\boldsymbol{\pi} - \mathbf{s})\mathbf{A}^{-1}||_1 = \sum_{i=1}^{n} |\pi_i - s_i|$ since $s_i(\alpha)$ changes from $s_i(0) = \pi_i$ to $s_i(1) = s_i$. For $||(\boldsymbol{\pi} - \mathbf{s})\mathbf{A}^{-1}||_\infty$, in the case where $|\pi_1 - s_1| \geq |\pi_i - s_i|,\ i = 2, 3, 4, \ldots, n$ and assuming uniform $\mathbf{s}$, monotonicity can be observed from changes in $\mathbf{s}\mathbf{A}^{-1}$ from $\mathbf{s}\mathbf{e}^T\boldsymbol{\pi} = \boldsymbol{\pi}$ to $\mathbf{s}\mathbf{I} = \mathbf{s}$.

Appendix B. Theoretical Bounds for One Norm $||\boldsymbol{\pi} - \boldsymbol{\psi}_{\boldsymbol{\alpha}}(\mathbf{s})||_1$

Appendix B.1. General Bound for Coevolutionary Digraphs

General bounds can be obtained for any (reducible and irreducible) coevolutionary digraph. A simple argument would be to take the maximum one norm for probability vectors $(0, 0, 0, \ldots, 1)$ and $(1, 0, 0, \ldots, 0)$ as 2 [23]. However, we can improve the bound by considering random walks on labelled or isomorphic tournaments [32, 33]. A probability vector $(a_1, a_2, a_3, \ldots, a_n)$ associated for a tournament with vertices consistently indexed according to their score sequences will have for its unilateral and directional dual (vertices indexed in the reversed order of their score sequences), $(a_n, a_{n-1}, a_{n-2}, \ldots, a_1)$ [34, 17].

Here, a straightforward digraph-theoretic argument would be to take the one norm between stationary vectors associated with transitive ($\boldsymbol{\pi}_{\text{tran}} = (0, 0, 0, \ldots, 1)$) and regular ($\boldsymbol{\pi}_{\text{reg}} = (\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n})$) tournaments

$$\begin{aligned} \max ||\star||_1 &= ||\boldsymbol{\pi}_{\text{tran}} - \boldsymbol{\pi}_{\text{reg}}||_1 \\ &= \Big(n-1\Big)\frac{1}{n} + \Big(1 - \frac{1}{n}\Big) \\ &= 2\Big(1 - \frac{1}{n}\Big) \end{aligned}$$

as the upper bound, that is, for the one norm of probability vectors associated with CMCs operating on tournaments $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$

$$||\boldsymbol{\pi}_1 - \boldsymbol{\pi}_2||_1 \leq 2\Big(1 - \frac{1}{n}\Big). \tag{A.8}$$

Appendix B.2. Bounds based on Digraph-Theoretic and Coupling Arguments

The bound given by the inequality in Equation A.8 also reflects the difference for a PageRank random walker operating on a transitive tournament between two opposite cases where $\alpha = 0$ so that all the probability mass is concentrated on a single absorbing vertex and where $\alpha = 1$, i.e. restarts all the time with a uniform teleportation vector so that the mass is now uniformly distributed on all vertices. What happens when α is varied in $(0, 1)$? In general, one would obtain loose bounds. In [35], a coupling argument is introduced to provide the relevant bounds, which we improve here.

Let $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ be the probability distributions of the random variables X and Y, respectively. Both X and Y take on values in the state space $\mathcal{V}$. Let (X, Y) be the pair of the two random variables. Let q be the joint distribution of (X, Y) on $\mathcal{V}\times\mathcal{V}$, i.e., $q(x, y) = \mathbb{P}(X = x, Y = y)$, such that $\sum_{y\in\mathcal{V}} q(x, y) = \mathbb{P}(X = x) = \mu_x$ and $\sum_{x\in\mathcal{V}} q(x, y) = \mathbb{P}(Y = y) = \nu_y$. A coupling of $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ refers to the pair (X, Y) defined on a single probability space whereby its marginal distributions of X and Y are $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$, respectively, i.e., satisfying $\mathbb{P}(X = x) = \mu_x$ and $\mathbb{P}(Y = y) = \nu_y$ for all $x, y \in \mathcal{V}$ [45].

The following result show how to bound the distance between stationary $\boldsymbol{\mu} = \boldsymbol{\pi}$ and PageRank $\boldsymbol{\nu} = \boldsymbol{\psi}_\alpha(\mathbf{s})$ distributions of the coupled CMCs $\{(X_t, Y_t) : t \in \mathbb{N}_0\}$, where $\{X_t : t \in \mathbb{N}_0\}$ is without restart and $\{Y_t : t \in \mathbb{N}_0\}$ is with restart. Our proof uses a combination of digraph-theoretic and coupling arguments. We exploit specific structures in irreducible coevolutionary digraphs $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$ to construct a specific coupling of the two CMCs. Then, we use of the equivalent characterizations of variation distance between the marginal distributions of the coupling ($||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_{\mathrm{TV}} = \frac{1}{2}||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1 = \min \mathbb{P}(X \neq Y)$) [46] and subsequently bound $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1$ with an asymptotic upper-bound of $\mathbb{P}(X_\infty \neq Y_\infty)$.

Lemma 7. *Let $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$ be an irreducible coevolutionary digraph. Let $\{(X_t, Y_t) : t \in \mathbb{N}_0\}$ be a coupling of the two CMCs. Associated with CMC $\{X_t : t \in \mathbb{N}_0\}$ is the probability transition matrix $\boldsymbol{P}$ and stationary* *distribution* $\boldsymbol{\pi}$*. Associated with CMC* $\{Y_t : t \in \mathbb{N}_0\}$ *is the perturbed probability transition matrix* $\boldsymbol{P}_{PR}$ *and stationary distribution* $\boldsymbol{\psi}_\alpha(\mathbf{s})$*. The lazy random walk versions are considered for both CMCs. The two stationary distributions satisfy the relation:*

$$\begin{aligned}
||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1 &\leq \frac{2}{1-\alpha}\mathbb{P}(X_\infty \neq Y_\infty) \\
&\leq \frac{2}{1-\alpha} d_\alpha \\
&\leq \frac{2}{1-\alpha}
\end{aligned}$$

where $d_\alpha = \max \mathbb{P}(X_{t+1} \neq Y_{t+1},\ X_t = Y_t \mid \text{restart at } t+1)$.

For $\{\tilde{Y}_t : t \in \mathbb{N}_0\}$ *with no restart* ($\alpha = 0$)*, there is an optimal coupling* $\{(X_t, \tilde{Y}_t) : t \in \mathbb{N}_0\}$ *such that* $||\boldsymbol{\pi} - \boldsymbol{\psi}_0(\mathbf{s})||_1 = 0$*, which implies* $\boldsymbol{\psi}_0(\mathbf{s}) = \boldsymbol{\pi}$*.*

Proof: Let $\{(X_t, Y_t) : t \in \mathbb{N}_0\}$ be a coupling of the two CMCs with the following construction.

(i) $X_0 = Y_0$ is drawn randomly from $\boldsymbol{\pi}$.

(ii) The state transitions are as follows: At time step $t+1$, decide with probability $1-\alpha$ not to restart or with probability α to restart.

  a. If there is no restart, then $X_{t+1} = Y_{t+1}$ always. If $X_t = Y_t$, then $X_{t+1} = Y_{t+1}$ (i.e., both chains jump to the same vertex) that is chosen randomly according to $\boldsymbol{P}$. If $X_t \neq Y_t$, then X_{t+1} and Y_{t+1} are chosen such that $X_{t+1} = Y_{t+1}$.

  b. If there is restart, then $X_{t+1} \neq Y_{t+1}$ always if $X_t \neq Y_t$. Otherwise, if $X_t = Y_t$, then X_{t+1} is chosen randomly according to $\boldsymbol{P}$ and Y_{t+1} is chosen uniformly at random such that $X_{t+1} \neq Y_{t+1}$.

For the coupling of the two CMCs, we first show that when there is no restart, $X_{t+1} = Y_{t+1}$ regardless of the state of the coupled CMCs at time t. Applying Moon's Theorem [32], an irreducible coevolutionary digraph $D_C(V_{S(n)}, A) \in \mathcal{D}_C(V_{S(n)})$ is vertex-pancyclic. In particular, any three distinct vertices $x, y, z \in V_{S(n)}$ forms a 3-cycle, i.e., cycle of length three. We use the simplified notation $xyzx$ to indicate $x \to y \to z \to x$. Let $X_t = x$ and $Y_t = y$. For lazy random walks, if the 3-cycle is $xzyx$, then arbitrarily set $X_{t+1} = X_t = x$ and $Y_{t+1} = x$. Similarly, if the 3-cycle is $yzxy$, then arbitrarily set $Y_{t+1} = Y_t = y$ and $X_{t+1} = y$. If X_t and Y_t are at the same vertex y, depending on the 3-cycle configuration, then both chains either jump to the same vertex z, i.e., $X_{t+1} = Y_{t+1} = z$ (if the 3-cycle is $xyzx$) or stay at vertex y (if the 3-cycle is $yxzy$). Similar arguments can be made for the case of $X_t = Y_t = x$.

When there is restart and if $X_t = Y_t$, we choose X_{t+1} randomly according to $\boldsymbol{P}$ and then choose Y_{t+1} uniformly at random such that $X_{t+1} \neq Y_{t+1}$. By construction, $X_{t+1} \neq Y_{t+1}$ always when $X_t \neq Y_t$. Although the two transitions are correlated, each of the two CMCs are still using their state transitions independently. As such, we have a coupled CMCs $\{(X_t, Y_t) : t \in \mathbb{N}_0\}$ such that their marginal distributions are $\boldsymbol{\pi}$ and $\boldsymbol{\psi}_\alpha(\mathbf{s})$.

We now use the coupling argument (Theorem 3 in [35]) for our coupling construction. Let $d_t = \mathbb{P}(X_t \neq Y_t)$. Since $X_0 = Y_0$, $\mathbb{P}(X_0 \neq Y_0) = 0$. Then,

$$\begin{aligned}
d_{t+1} &= \mathbb{P}(X_{t+1} \neq Y_{t+1}) \\
&= \mathbb{P}(X_{t+1} \neq Y_{t+1} \mid \text{no restart at } t+1)\mathbb{P}(\text{no restart}) \\
&\quad + \ \mathbb{P}(X_{t+1} \neq Y_{t+1} \mid \text{restart at } t+1)\mathbb{P}(\text{restart}) \\
&= (0)(1-\alpha) + \alpha\big(\mathbb{P}(X_{t+1} \neq Y_{t+1},\ X_t \neq Y_t \mid \text{restart at } t+1)\big) \\
&\quad + \ \alpha\big(\mathbb{P}(X_{t+1} \neq Y_{t+1},\ X_t = Y_t \mid \text{restart at } t+1)\big) \\
&\leq \alpha\big(\mathbb{P}(X_t \neq Y_t) + \mathbb{P}(X_{t+1} \neq Y_{t+1},\ X_t = Y_t \mid \text{restart at } t+1)\big) \\
&= \alpha\big(d_t + \mathbb{P}(X_{t+1} \neq Y_{t+1},\ X_t = Y_t \mid \text{restart at } t+1)\big).
\end{aligned}$$

Since we start with $d_0 = 0$ and taking $d_\alpha = \max \mathbb{P}(X_{t+1} \neq Y_{t+1},\ X_t = Y_t \mid \text{restart at } t+1)$ with a slight abuse of notation, as we iterate the bound on $d_{t+1} \leq \alpha(d_t + d_\alpha)$, we obtain $d_1 = (\alpha)d_\alpha$, $d_2 = (\alpha + \alpha^2)d_\alpha$, $d_3 = (\alpha + \alpha^2 + \alpha^3)d_\alpha$, $\ldots$, which follows a geometric progression. We can asymptotically bound $d_\infty \leq \frac{1}{1-\alpha}d_\alpha$. Furthermore, since (X_∞, Y_∞) is drawn from the stationary distribution of the correlated chains, the marginal distributions of X_∞ and Y_∞ are $\boldsymbol{\pi}$ and $\boldsymbol{\psi}_\alpha(\mathbf{s})$, respectively. As such, $\mathbb{P}(X_\infty \neq Y_\infty) = d_\infty \leq \frac{1}{1-\alpha}d_\alpha$.

We apply the *Coupling Lemma* (Lemma 2.19 in [46]) to obtain the variation distance between the two marginal distributions

$$||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_{\text{TV}} = \frac{1}{2}||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1$$
$$= \min \mathbb{P}(X \neq Y)$$

to obtain the bound $||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1 \leq 2\mathbb{P}(X_\infty \neq Y_\infty) = 2d_\infty \leq \frac{2}{1-\alpha}d_\alpha \leq \frac{2}{1-\alpha}$.

Where there is no restart, the CMC is a copy of $\{X_t : t \in \mathbb{N}_0\}$. Let $\{\tilde{Y}_t : t \in \mathbb{N}_0\}$ denote such a CMC. Starting with $X_0 = \tilde{Y}_0$ that is drawn randomly from $\boldsymbol{\pi}$, there is an optimal coupling $\{(X_t, \tilde{Y}_t) : t \in \mathbb{N}_0\}$ whereby $X_{t+1} = \tilde{Y}_{t+1}$ always. In this case, $X_0 = \tilde{Y}_0, X_1 = \tilde{Y}_1, X_2 = \tilde{Y}_2, \ldots$ whereby $||\boldsymbol{\pi} - \boldsymbol{\psi}_0(\mathbf{s})||_{\text{TV}} = \min \mathbb{P}(X_\infty, Y_\infty) = 0$, which implies that $\boldsymbol{\psi}_0(\mathbf{s}) = \boldsymbol{\pi}$. □

Corollary 8. *Let $\mathcal{D}_C(V_{S(n)})$ be the set of irreducible coevolutionary digraphs. Let any $D_C(V_{S(n)}) \in \mathcal{D}_C(V_{S(n)})$ be an irreducible coevolutionary digraph so that the vertices $v_1, v_2, v_3, \ldots, v_n$ are ordered according to the score sequence $d^-_T(v_1) \leq d^-_T(v_2) \leq d^-_T(v_3) \leq \cdots \leq d^-_T(v_n)$, where $d^-_T(v_i)$ is the in-degree of vertex v_i. Assume the digraph having the following structure:*

1. *The set of vertices $V_{S(n)}$ can be partitioned into three disjoint subsets with $\{v_1\} \cup V_{S(n-2)} \cup \{v_n\} = V_{S(n)}$, where $V_{S(n-2)} = V_{S(n)} \backslash \{v_1, v_n\}$, $\{v_1\} \cap V^m = \emptyset$, and $\{v_n\} \cap V_{S(n-2)} = \emptyset$.*
2. *$\{v_1\} \mapsto V_{S(n-2)} \mapsto \{v_n\}$. That is, $v_1 \to v_i$ for all $v_i \in V_{S(n-2)}$ and that $v_i \to v_n$ for all $v_i \in V_{S(n-2)}$.*
3. *$v_1 \to v_i \to v_n \to v_1$ forms a 3-cycle for all $v_i \in V_{S(n-2)}$.*

For a coupled CMCs $\{(X_t, Y_t) : t \in \mathbb{N}_0\}$ operating on $D_C(V_{S(n)})$, there is the following bound in distance

$$||\boldsymbol{\pi} - \boldsymbol{\psi}_\alpha(\mathbf{s})||_1 \leq \frac{2}{1-\alpha}\left(1 - \left(\pi_{v_n} - \frac{1}{n}\right)\right).$$

Proof: Since $X_0 = Y_0$ is drawn randomly from $\boldsymbol{\pi}$ and that (X_∞, Y_∞) is drawn from the stationary distribution of the correlated chains, we can bound d_∞ as follows

$$d_\infty \leq \frac{1}{1-\alpha}\mathbb{P}(X_\infty \neq Y_\infty)$$

$$
\begin{aligned}
&\le \frac{1}{1-\alpha}\big(1-\mathbb{P}(X_\infty = Y_\infty)\big) \\
&\le \frac{1}{1-\alpha}\left(1-\sum_{v_i \in V_{S(n)}} \mathbb{P}\big((X_\infty = v_i)\wedge (Y_\infty = v_i)\big)\right) \\
&\le \frac{1}{1-\alpha}\left(1-\sum_{v_i \in V_{S(n)}} \min\{\mathbb{P}(X_\infty = v_i), \mathbb{P}(Y_\infty = v_i)\}\right) \\
&\le \frac{1}{1-\alpha}\Big(1-\min\{\mathbb{P}(X_\infty = v_n), \mathbb{P}(Y_\infty = v_n)\}\Big) \\
&\le \frac{1}{1-\alpha}\left(1-\left(\pi_{v_n} - \frac{1}{n}\right)\right).
\end{aligned}
$$

Our argument proceeds as follows. By construction, $(X_\infty = Y_\infty)$ occurs when there is no restart. (X_∞, Y_∞) is drawn from the stationary distribution of the correlated chains with the marginal distributions of X_∞ and Y_∞ are $\boldsymbol{\pi}$ and $\boldsymbol{\psi}_\alpha(\mathbf{s})$, respectively. We can use the bound $\mathbb{P}\big((X_\infty = v_i)\wedge(Y_\infty = v_i)\big) \le \min\{\mathbb{P}(X_\infty = v_i), \mathbb{P}(Y_\infty = v_i)\}$ for the third inequality [45]. In this manner, $\mathbb{P}(X_\infty = v_i) = \pi_{v_i}$ and $\mathbb{P}(Y_\infty = v_i) = \psi_{v_i}$ for all $v_i \in V_{S(n)}$. For the fifth inequality, given that probabilities range in $[0, 1]$, we could choose the vertex that the CMC would most often jump to (i.e., v_n) as the upper bound. For the sixth inequality, $\min\{\pi_{v_n}, \psi_{v_n}\} = \psi_{v_n}$ but we can bound the term $1 - \min\{\pi_{v_n}, \psi_{v_n}\}$ by noting that $\pi_{v_n} - \frac{1}{n} \le \psi_{v_n} \le \pi_{v_n}$.

Since X_∞ is drawn from the stationary distribution $\boldsymbol{\pi}$, we need to show that $\pi_{v_i} \le \pi_{v_1} \le \pi_{v_n}$ for all $v_i \in V_{S(n-2)}$. We use a digraph-theoretic argument. We can assume that the subdigraph $D_C(V_{S(n-2)})$ induced by $V_{S(n-2)}$ is transitive. We apply the *Canonical Decomposition of quasi-transitive digraphs* (Theorem 4.8.5 [17]) on $D_C(V_{S(n)}, A)$ and obtain a strong semicomplete digraph (in this case, a digraph of 3-cycle) $u_1 \to u_2 \to u_3 \to u_1$ whereby $u_1 = v_1$, $u_3 = v_n$, and u_2 is the *contraction* of $V_{S(n-2)}$. We can then view the CMC as a random walk that is jumping clockwise along the direction of $u_1 \to u_2 \to u_3 \to u_1$.

Given that the score sequence $d_T^-(v_1) \le d_T^-(v_2) \le d_T^-(v_3) \le \cdots \le d_T^-(v_n)$ of $D_C(V_{S(n)})$ orders the vertices according to its in-degrees, a random walk on this digraph will spend most of the fraction of its time at v_n. Note that although $d_T^-(v_n) = d_T^-(v_{n-1})$ in the case $D_C(V_{S(n-2)})$ is transitive, v_{n-1} is oriented towards v_n. Furthermore, although for any $v_i \in V_{S(n-2)} \backslash \{v_{n-1}\}$, there could be equal probability (for a *standard* random walk) of jumping towards v_{n-1} and v_n, the relation $v_{n-1} \rightarrow v_n$ ensures that any time the random walk jumps to v_{n-1} it must then jump to v_n. At other times, the random walk could jump directly to v_n so that on average the random walk spends more fraction of its time in v_n. We also note that since v_n is oriented towards v_1 only, every time the random walk jumps to v_n it will then jump to v_1, in which case $\pi_{v_1} = \pi_{v_n}$. As such, $\pi_{v_i} \leq \pi_{v_1} = \pi_{v_n}$ for all $v_i \in V_{S(n-2)}$. $\square$

Note: For an irreducible coevolutionary digraph $D_C(V_{S(n)})$ having structures described in Corollary 8, it is possible to compute π_{v_n} directly without having to compute the full $\boldsymbol{\pi}$. When $D_C(V_{S(n)})$ is pancyclic with the least number of 3-cycles, we can calculate $\pi_{v_n} = \pi_{v_1} = \frac{1}{2}(1 - \sum_{i=2}^{n-1} \pi_i)$ using only hitting times. Note $D_C(V_{S(n-2)})$ is transitive. We calculate $1 - \sum_{i=2}^{n-1} \pi_i$ as the fraction of time on average the random walk spends jumping along the subdigraph $D_C(V_{S(n-2)})$ as it cycles along $u_1 \rightarrow u_2 \rightarrow u_3 \rightarrow u_1$. More precisely, it is the average of the expected hitting time to v_n over starting vertices $v_i \in V_{S(n-2)}$.

The subdigraph induced by $V_{S(n-2)} \cup \{v_n\}$ is transitive and that $\{v_1\} \mapsto V_{S(n-2)}$. So, we can treat the random walk on the subdigraph as an absorbing CMC with transient states $v_i \in V_{S(n-2)}$ and a single absorbing state v_n. Let $\mathbb{E}_{v_i}(\tau_{v_n})$ be the expected hitting time starting from v_i to v_n. We relabel the vertices as x_m, $m = n - i$, $i = 1, 2, 3, \ldots, n-2$ (e.g., vertex v_{n-1} is now x_1). It can be shown that $\mathbb{E}_{v_i}(\tau_{v_n}) = \mathbb{E}_{x_m}(\tau_{v_n}) = \sum_{j=1}^{m} \frac{1}{j}$. In this case, we have

$$
\begin{aligned}
\mathbb{E}_{v_{n-1}}(\tau_{v_n}) &= \mathbb{E}_{x_1}(\tau_{v_n}) = 1 \\
\mathbb{E}_{v_{n-2}}(\tau_{v_n}) &= \mathbb{E}_{x_2}(\tau_{v_n}) = 1 + \frac{1}{2} \\
&\vdots \\
\mathbb{E}_{v_2}(\tau_{v_n}) &= \mathbb{E}_{x_{n-2}}(\tau_{v_n}) = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n-2}.
\end{aligned}
$$

Let $\mathbb{E}(\eta_{V_{S(n-2)}}) = \frac{1}{n-2} \sum_{m=1}^{n-2} \mathbb{E}_{x_m}(\tau_{v_n})$ be the fraction of time on average the random walk spends in $V_{S(n-2)}$ is as it cycles through $u_1 \rightarrow u_2 \rightarrow u_3 \rightarrow u_1$. $\mathbb{E}(\eta_{v_1}) = 1$ and $\mathbb{E}(\eta_{v_n}) = 1$. So the total average time would be $2 + \mathbb{E}(\eta_{V_{S(n-2)}})$.

We can calculate $1 - \sum_{i=2}^{n-1} \pi_i$ using these fraction of times as the random walk cycles through $u_1 \to u_2 \to u_3 \to u_1$ as follows

$$\pi_{v_n} = \pi_{v_1} = \frac{1}{2}\left(1 - \frac{\mathbb{E}(\eta_{V_{S(n-2)}})}{2 + \mathbb{E}(\eta_{V_{S(n-2)}})}\right). \tag{A.9}$$

References

[1] R. Axelrod, The evolution of strategies in the iterated prisoner's dilemma, in: L. D. Davis (Ed.), Genetic Algorithms and Simulated Annealing, Morgan Kaufmann, New York, 1987, Ch. 3, pp. 32–41.

[2] K. Sims, Evolving 3D morphology and behavior by competition, in: R. Brooks, P. Maes (Eds.), Artificial Life IV, MIT Press, Cambridge, MA, 1994, pp. 28–39.

[3] D. B. Fogel, Blondie24: Playing at the Edge of AI, Morgan Kaufmann, San Francisco, CA, 2002.

[4] G. Kendall, X. Yao, S. Y. Chong, The Iterated Prisoners' Dilemma: 20 Years On, World Scientific, Singapore, 2007.

[5] K. Chellapilla, D. B. Fogel, Evolution, neural networks, games, and intelligence, Proceedings of the IEEE 87 (9) (1999) 1471–1496.

[6] D. Cliff, G. F. Miller, Tracking the red queen: Measurements of adaptive progress in co-evolutionary simulations, in: Advances in Artificial Life: Proc. of the Third European Conference on Artificial Life, Vol. 929 of Lecture Notes in Computer Science, Springer-Verlag, 1995, pp. 200–218.

[7] K. O. Stanley, R. Miikkulainen, The dominance tournament method of monitoring progress in coevolution, in: Bird of a Feather Workshop: Proc. of the 2002 Genetic and Evolutionary Computation Conference (GECCO'02), New York, USA, 2002, pp. 242–248.

[8] R. Janssen, S. Nolfi, P. Haselager, I. Sprinkhuizen-Kuyper, Cyclic incrementality in competitive coevolution: Evolvability through pseudo-baldwinian switching-genes, Artificial Life 22 (3) (2016) 319–352.

[9] S. G. Ficici, O. Melnik, J. B. Pollack, A game-theoretic and dynamical-systems analysis of selection methods in coevolution, IEEE Transactions on Evolutionary Computation 9 (6) (2005) 580–602.

[10] P. Tiňo, S. Y. Chong, X. Yao, Complex co-evolutionary dynamics - structural stability and finite population effects, IEEE Transactions on Evolutionary Computation 17 (2) (2013) 155–164.

[11] S. Samothrakis, S. Lucas, T. P. Runarsson, Coevolving game-playing agents: Measuring performance and intransitivities, IEEE Transactions on Evolutionary Computation 17 (2) (2013) 213–226.

[12] S. Y. Chong, P. Tiňo, J. He, X. Yao, A new framework for analysis of coevolutionary systems - directed graph representation and random walks, Evolutionary Computationdoi:`https://doi.org/10.1162/evco{\_}a{\_}00218`.

[13] J. He, X. Yao, Towards an analytic framework for analysing the computation time of evolutionary algorithms, Artificial Intelligence 145 (2003) 59–97.

[14] D. F. Gleich, Pagerank beyond the web, SIAM Review 57 (3) (2015) 321–363.

[15] S. Brin, L. Page, The anatomy of a large-scale hypertextual web search engine, in: Proc. of the 7th International World Wide Web Conference (WWW7), Brisbane, Australia, 1998, pp. 107–117.

[16] J. M. Kleinberg, Authoritative sources in a hyperlinked environment, Journal of the ACM 46 (5) (1999) 604–632.

[17] J. Bang-Jensen, G. Z. Gutin, Digraphs Theory, Algorithms and Applications, Springer-Verlag, London, 2009.

[18] T. P. Runarsson, S. M. Lucas, Coevolution versus self-play temporal difference learning for acquiring position evaluation in small-board go, IEEE Transactions on Evolutionary Computation 9 (6) (2005) 540–551.

[19] J. R. Norris, Markov Chains, Cambridge University Press, Cambridge, UK, 1998.

[20] S. Meyn, R. L. Tweedie, Markov Chains and Stochastic Stability, 2nd Edition, Cambridge University Press, Cambridge, UK, 2009.

[21] M. Bianchini, M. Gori, F. Scarselli, Inside pagerank, ACM Transactions on Internet Technology 5 (1) (2005) 92–128.

[22] C. D. Meyer, Matrix Analysis and Applied Linear Algebra, SIAM, Philadelphia, PA, 2000.

[23] A. N. Langville, C. D. Meyer, Inside pagerank, ACM Transactions on Internet Technology 1 (3) (2004) 335–380.

[24] G. Jeh, J. Widom, Scaling personalized web search, in: Proc. of the 12th World Wide Web Conference (WWW'03), Budapest, Hungary, 2003, pp. 271–279.

[25] F. Chung, W. Zhao, Pagerank and random walks on graphs, in: Fete of Combinatorics and Computer Science, Vol. 20 of Bolyai Society Mathematical Studies, Springer, 2010, pp. 43–62.

[26] R. A. Horn, C. R. Johnson, Matrix Analysis, 2nd Edition, Cambridge University Press, Cambridge, UK, 2012.

[27] A. Berman, R. J. Plemmons, Nonnegative Matrices in the Mathematical Sciences, SIAM, Philadelphia, PA, 1994.

[28] G. H. Golub, C. F. Van Loan, Matrix Computation, 4th Edition, The Johns Hopkins University Press, Maryland, USA, 2013.

[29] Y. Li, Z. Zhang, Random walks on digraphs, the generalized digraph laplacian and the degree of asymmetry, in: Algorithms and Models for the Web-Graph (WAW 2010), Vol. 6516 of Lecture Notes in Computer Science, Springer, 2010, pp. 74–85.

[30] H. V. Henderson, S. R. Searle, On deriving the inverse of a sum of matrices, SIAM Review 23 (1) (1981) 53–60.

[31] P. J. Schweitzer, Perturbation theory and finite markov chains, Journal of Applied Probability 5 (2) (1968) 401–413.

[32] J. W. Moon, Topics on Tournament, Holt, Rinehart and Winston, New York, 1968.

[33] P. Tetali, A characterization of unique tournaments, Journal of Combinatorial Theory, Series B 72 (1998) 157–159.

[34] T. Burns, L. D. Meeker, A mathematical model of evaluation, decision-making, and social interaction, in: J. Cochrane, M. Zeleny (Eds.), Multiple Criteria Decision-Making, University of South Carolina Press, Columbia, South Carolina, 1973, pp. 141–163.

[35] A. Y. Ng, A. X. Zheng, M. I. Jordan, Link analysis, eigenvectors and stability, in: Proc. of the 17th International Joint Conference on Artificial Intelligence (IJCAI'01), Seattle, USA, 2001, pp. 903–910.

[36] T. H. Haveliwala, Topic-sensitive pagerank: A context-sensitive ranking algorithm for web search, IEEE Transactions on Knowledge and Data Engineering 15 (4) (2003) 784–796.

[37] M. Burzio, J. Pelant, On complete strongly connected digraphs with the least number of 3-cycles, Discrete Mathematics 155 (1996) 27–30.

[38] B. Rudolf, M. Markošová, M. Čajági, P. Tiňo, Degree distribution and scaling in the connecting-nearest-neighbors model, Physical Review E 85 (2) (2012) 026114.

[39] J. H. Conway, R. K. Guy, The Book of Numbers, Springer-Verlag, New York, 1996.

[40] M. G. Kendall, Rank Correlation Methods, 4th Edition, Charles Griffin, London, 1975.

[41] C. Dwork, R. Kuman, M. Naor, D. Sivakumar, Rank aggregation methods for the web, in: Proc. of the 10th International Joint Conference on World Wide Web (WWW'01), Hong Kong, 2001, pp. 613–622.

[42] S. Herrmann, G. Ochoa, F. Rothlauf, Pagerank centrality for performance prediction: The impact of the local optima network model, Journal of Heuristicsdoi:https://doi.org/10.1007/s10732-017-9333-1.

[43] N. Morača, Bounds for norms of the matrix inverse and the smallest singular value, Linear Algebra and its Applications 429 (2008) 2589–2601.

[44] C. R. Johnson, Inverse m-matrices, Linear Algebra and its Applications 47 (1982) 195–216.

[45] D. A. Levin, Y. Peres, E. L. Wilmer, Markov Chains and Mixing Times, American Mathematical Society, Providence, Rhode Island, 2009.

[46] D. J. Aldous, J. A. Fill, Reversible markov chains and random walks on graphs, unfinished monograph, recompiled 2014, available at http://www.stat.berkeley.edu/∼aldous/RWG/book.html (1983).