*A new framework for analysis of coevolutionary systems*
Chong, Siang Yew; Tiňo, Peter; He, Jun; Yao, Xin

# A New Framework for Analysis of Coevolutionary Systems - Directed Graph Representation and Random Walks

**Siang Yew Chong**                                    S.Y.Chong@cs.bham.ac.uk

School of Computer Science, University of Birmingham, Birmingham B15 2TT, United Kingdom and School of Computer Science, University of Nottingham, Malaysia Campus, Jalan Broga, 43500 Semenyih, Malaysia

**Peter Tiňo**                                         P.Tino@cs.bham.ac.uk

School of Computer Science, University of Birmingham, Birmingham B15 2TT, United Kingdom

**Jun He**                                             Jun.He@aber.ac.uk

Department of Computer Science, Aberystwyth University, Aberystwyth SY23 3DB, United Kingdom

**Xin Yao**                                            X.Yao@cs.bham.ac.uk

School of Computer Science, University of Birmingham, Birmingham B15 2TT, United Kingdom

**Abstract**

Studying coevolutionary systems in the context of simplified models (i.e. games with pairwise interactions between coevolving solutions modelled as self plays) remains an open challenge since the rich underlying structures associated with pairwise-comparison-based fitness measures are often not taken fully into account. Although cyclic dynamics have been demonstrated in several contexts (such as intransitivity in coevolutionary problems), there is no complete characterization of cycle structures and their effects on coevolutionary search. We develop a new framework to address this issue. At the core of our approach is the directed graph (digraph) representation of coevolutionary problem that fully captures structures in the relations between candidate solutions. Coevolutionary processes are modelled as a specific type of Markov chains – random walks on digraphs. Using this framework, we show that coevolutionary problems admit a qualitative characterization: a coevolutionary problem is either solvable (there is a subset of solutions that dominates the remaining candidate solutions) or not. This has an implication on coevolutionary search. We further develop our framework that provide the means to construct quantitative tools for analysis of coevolutionary processes and demonstrate their applications through case studies. We show that coevolution of solvable problems corresponds to an absorbing Markov chain for which we can compute the expected hitting time of the absorbing class. Otherwise, coevolution will cycle indefinitely and the quantity of interest will be the limiting invariant distribution of the Markov chain. We also provide an index for characterizing complexity in coevolutionary problems and show how they can be generated in a controlled manner.

**Keywords**

Coevolution, Directed graphs, Markov chains, Evolutionary Computation, Self play.

# 1 Introduction

In this study we focus on coevolutionary systems that are specified by the: (1) *coevolutionary problem* – the set of all alternatives (e.g. strategies in a two-player game) and a pairwise preference relation indicating the superior alternative (e.g. whether when playing the game one strategy defeats another one), and (2) *coevolutionary process* – the population of alternatives undergoing a process of selection and variation that is guided only by their interaction outcomes. Coevolutionary systems are often used in competitive settings where the lack of absolute fitness measures necessitates that one exploit the strategic nature of interactions in coevolving populations. In this context, coevolutionary systems provide a natural alternative for solving test-based problems (Hillis, 1990; de Jong and Pollack, 2004) that include two-player strategic games whereby opponent strategies provide test cases for fitness evaluations (Runarsson and Lucas, 2005). In turn, game-theoretic-based frameworks have been developed to provide tools for analysis of coevolutionary systems (Ficici et al., 2005; Tiňo et al., 2013).

The use of relative and dynamic fitness evaluations in coevolutionary systems introduces unique challenges that make these systems difficult to analyze and understand in comparison to evolutionary systems. They are collectively referred as coevolutionary pathologies. In the classic example of two-population predator-prey coevolution, the *Red Queen effect* is coined to describe the interdependence in the fitness evaluations of members in one population to the other even though there is no exchange of genetic materials (Cliff and Miller, 1995; Janssen et al., 2016). *Disengagement* occurs when there are large gaps in competence level between the populations, leading to a loss of selection pressure, decoupling of the coevolutionary system, and then onset of random drift in the populations (Cartlidge and Bullock, 2004). For problems with intransitivity in the relationships between the solutions, *cyclic dynamics* have been observed whereby the population overspecializes to copies of a solution that is exploited by another solution later in the coevolutionary process (Darwen and Yao, 1997). In the more extreme case, coevolution *forgets* a solution it has discovered earlier – the solution that is driven to extinction is adaptively found later (Rosin and Belew, 1997).

Although there are broader implications in design of effective coevolutionary systems, pathologies are usually associated with the selection part of the coevolutionary process particularly on the use of pairwise-comparison-based fitness measures. Most studies would frame the pathologies in specific settings for which relevant methodologies can be used for analysis. These include measurements for monitoring progress of coevolutionary systems using generational samples of test cases (Cliff and Miller, 1995; Stanley and Miikkulainen, 2002; Bader-Natal and Pollack, 2005) and those where one has a global view and can objectively measure performance of solutions (Floreano and Nolfi, 1997; Chong et al., 2012). In (Ficici et al., 2005; Tiňo et al., 2013), a formal approach is taken to study coevolutionary dynamics. Under the usual evolutionary game theoretic setting such as the use of an infinite population, a full characterization is given on coevolutionary dynamics (as changes to the ratio of individuals adopting a pure strategy) for various dynamical maps modelling different selection operations.

Here, we focus on coevolutionary systems that use pairwise-comparison-based fitness measures. However, it remains an open challenge to study this restricted class of coevolutionary systems since the rich underlying structures associated with pairwise-comparison-based fitness measures are often not taken fully into account. Although cyclic dynamics have been shown in coevolutionary systems, most past studies such as (Samothrakis et al., 2013) would focus on short cycles (of length three). Longer length cycles can occur in coevolution. Crucially, there is no complete characterization of cycle

structures and their effects on coevolution. Our motivation is to develop a new framework to address this issue formally. Our approach uses the digraph representation for coevolutionary problems to capture the full structures in the binary relations between candidate solutions. Coevolutionary processes are modelled naturally as discrete time Markov chains operating on discrete state spaces (the sets of vertices of digraphs). In this manner, different coevolutionary algorithms would correspond to different implementations of random walks on digraphs.

We first show that coevolutionary problems admit a qualitative characterization: a coevolutionary problem is either solvable (the digraph is reducible with a subset of solutions that dominates the remaining solutions of the solution space) or not. This has an implication on coevolutionary search where we further develop the framework to construct quantitative tools for appropriate analysis and demonstrate their applications through case studies. We will show that coevolution of solvable problems corresponds to an absorbing Markov chain for which we can compute the expected hitting time of the absorbing class. Otherwise, coevolution will cycle indefinitely and the quantity of interest will be the limiting invariant distribution of the Markov chain. We also suggest an index for characterizing complexity in coevolutionary problems and show how a hierarchy of increasingly complex coevolutionary problems can be generated in a controlled manner.

This paper is organized as follows. In Section 2, we formulate the digraph representation and provide several results characterizing cycle structures in coevolutionary problems. Section 3 formulates coevolutionary processes as random walks on digraphs and provides examples how specific Markov chains are constructed from coevolutionary algorithms operating on digraphs. Section 4 introduces relevant quantitative characterizations for coevolutionary systems (expected hitting times and stationary distributions of Markov chains). These quantitative tools will be demonstrated through theoretical and computational case studies. Section 5 will discuss the main findings and wider implications of our framework with remarks for future studies.

## 2 Digraph Representation of Coevolutionary Problem

We first present basic definitions and facts on digraphs relevant to formulating our framework to make this paper self-contained. Our notation and terminology follow standard digraph theory texts and readers can refer (Bang-Jensen and Gutin, 2009) for detailed expositions of the following. Our focus is to show how digraphs provide appropriate mathematical structures that fully capture the pairwise preference relations between objects of interest that one would encounter in typical coevolutionary problems such as games. Later, we would exploit classic results in digraph theory to provide a complete characterization of cycle structures in coevolutionary problems.

### 2.1 Digraphs - Definition, Notation and Terminology

**Definition 2.1.1.** A *digraph* $D = (V, A)$ consists of two sets: (i) a non-empty, finite set $V(D)$ of elements called *vertices*, and (ii) a finite set $A(D)$ of elements corresponding to ordered pairs $(u, v)$ of *distinct* vertices, $u, v \in V(D)$, called *arcs* (or *directed edges*). The arc set is a *binary relation* on the vertex set ($A$ is a subset of the *Cartesian product* of $V$ with itself, $A \subset V \times V$).

**Definition 2.1.2.** The *order* of a digraph $D = (V, A)$ is given by the *number of vertices* $|V|$ (often denoted as $|D|$). The *size* of $D$ is given by the *number of arcs* $|A|$.

**Definition 2.1.3.** A graph $G = (V, E)$ consists of: (i) a non-empty, finite vertex set $V(G)$, and (ii) a finite set $E(G)$ corresponding to unordered pairs $\{x, y\}$ of *distinct* vertices, $x, y \in V(G)$, called *edges*. The *underlying graph* $UG(D)$ of a digraph $D = (V, A)$

is the graph with the same vertex set $V$ and the corresponding edge set $E$ whereby arcs are replaced by edges with no parallel edges and loops.

We often use a simpler notation and denote the arc $(u, v)$ as $uv$. The tail $u$ and head $v$ are the *end-vertices* of $uv$. *Adjacency* between end-vertices $u$ and $v$ implies an arc $uv$ or $vu$. As in (Bang-Jensen and Gutin, 2009), we consider the digraph $D = (V, A)$ without *parallel* arcs between adjacent end-vertices (arcs having the same tail and same head) and *loops* (the arc $uu$ where the tail and head is the same vertex).[1]

The definition of digraphs and in particular the binary relation on the vertex set that can be constructed allows for rich structures to be captured. We will consider notions of orientation, biorientation, complete, and semicomplete.

**Definition 2.1.4.** The *orientation* of a graph is the digraph where each edge corresponding to a pair of adjacent vertices $u, v \in V$ is connected by an arc (either $uv$ or $vu$ only). The *biorientation* of a graph is the digraph where each edge in the graph is replaced by the arc $uv$ or $vu$ or both arcs $uv$ and $vu$.

Obviously, any digraph is a biorientation. An oriented graph is a specific biorientation since any pair of adjacent vertices is connected by an arc only. In a *complete* graph, every pair of distinct vertices is adjacent.

**Definition 2.1.5.** A *semicomplete digraph* is a biorientation of a complete graph (since not all pairs of distinct vertices are completely bioriented, in which case they are only oriented). A *tournament* is a complete oriented graph.

Tournaments form a special and well-studied class of semicomplete digraphs (Bang-Jensen and Gutin, 2009). Tournaments and their generalizations (semicomplete digraphs) would be the focus later in this paper.

## 2.2 Neighbourhoods and Degrees

**Definition 2.2.1.** A *neighbourhood* of a vertex $u \in V$, $N_D(u)$, refers to the subset of $V$ whereby its elements are adjacent to $u$. It consists of the union of the *out-neighbourhood* $N_D^+(u)$ and *in-neighbourhood* $N_D^-(u)$ of $u$:

$$N_D^+(u) = \{v \in V - \{u\} : uv \in A\}, \quad N_D^-(u) = \{w \in V - \{u\} : wu \in A\}.$$

We note $V - \{u\} \subset V$ results from the set *difference* between $V$ and $u$ in the usual manner. The vertices in $N_D^+(u)$, $N_D^-(u)$, and $N_D(u) = N_D^+(u) \cup N_D^-(u)$ are called *out-neighbours*, *in-neighbours*, and *neighbours*, respectively (Bang-Jensen and Gutin, 2009).

A related notion of neighbourhoods is the *degrees* in a digraph.

**Definition 2.2.2.** The sets of outgoing arcs $(u, V - u)_D$ and incoming arcs $(V - u, u)_D$ are defined, respectively, as

$$(u, V - \{u\})_D = \{(u, v) \in A : v \in V - \{u\}\}, \quad (V - \{u\}, u)_D = \{(w, u) \in A : w \in V - \{u\}\}.$$

The *out-degree* $d_D^+(u)$ (*in-degree* $d_D^-(u)$) of a vertex $u$ is the number of outgoing (incoming) arcs in the digraph $D$ where $u$ is the tail (head) for each of these arcs. They are given by $d_D^+(u) = |(u, V - \{u\})_D|$ and $d_D^-(u) = |(V - \{u\}, u)_D|$. The degree of $u$ is given by $d_D(u) = d_D^+(u) + d_D^-(u)$.

---

[1] We do consider loops later when we introduce Markov chains, which essentially are weighted digraphs with specific constraints.

## 2.3 Walks, Paths, and Cycles

Consider a digraph $D = (V, A)$, whereby vertices $v_i$ and arcs $a_i$ from $D$ are labelled in a manner such that the tail and head of $a_i$ is $v_i$ and $v_{i+1}$, respectively, for every $i = 1, 2, 3, \ldots, k-1$ (Bang-Jensen and Gutin, 2009).

**Definition 2.3.1.** A $(v_1, v_k)$-*walk* $\mathcal{W}$ in the digraph $D$ from $v_1$ to $v_k$ is the alternating sequence of vertices and arcs $v_1 a_1 v_2 a_2 v_3 \ldots a_{k-1} v_k$. The set of vertices of the walk is $V(\mathcal{W}) = \{v_1, \ldots, v_k\}$ while the set of arcs of the walk is $A(\mathcal{W}) = \{a_1, \ldots, a_{k-1}\}$. $\mathcal{W}$ is said to be *closed* if $v_1 = v_k$, otherwise it is *open*. The length of a walk is the number of arcs in it and by definition of $\mathcal{W}$ is simply $k$.

**Definition 2.3.2.** A $(v_1, v_k)$-*path* $\mathcal{P}$ is a walk where all the vertices in it are distinct. A *k-cycle* $\mathcal{C}$ is a closed $(v_1, v_k)$-walk of length $k \geq 3$ on the $(v_1, v_{k-1})$-path.[2]

We usually denote $\mathcal{W}$, $\mathcal{P}$, and $\mathcal{C}$ without the arcs, $v_1 v_2 v_3 \ldots v_k$, since $a_i$ is defined in the context of $v_i v_{i+1}$ (Bang-Jensen and Gutin, 2009). $\mathcal{P}$ and $\mathcal{C}$ on $D$ are by definition *subdigraphs* of $D$, e.g., $V(\mathcal{P}) \subseteq V(D)$, $A(\mathcal{P}) \subseteq A(D)$, and $A(\mathcal{P}) = \{uv : u, v \in V(\mathcal{P})\}$.

**Definition 2.3.3.** The path $\mathcal{P}$ is said to be *Hamilton* if $V(\mathcal{P})$ is the same as $V(D)$. A Hamilton cycle $\mathcal{C}$ has $V(\mathcal{C}) = V(D)$. A digraph $D$ is said to be *hamiltonian* if it contains a Hamilton cycle.

## 2.4 Classes of Coevolutionary Problems as Digraphs

We state how coevolutionary problems can be represented as digraphs in the following.

**Definition 2.4.1.** A coevolutionary problem is defined as an ordered pair $C = (S, R)$ where $S$ represents the solution set and $R$ is the pairwise preference relation on $S$. This is represented as the digraph $D_C = (V_S, A_R)$. The vertex set is the solution set $V_S = S$ (each distinct vertex corresponds to a single unique solution). The arc set is the binary relation on $S$, $A_R = R$, (each biorientation corresponds to the pairwise preference relationship between each pair of solutions).

The underlying graph $UG(D_C)$ is complete since there is a relationship between every pair of solutions in a coevolutionary problem. Each problem is specified by the number of unique solutions and associated preference relation on those solutions, which is described by the biorientations on $UG(D_C)$. We use the term *coevolutionary digraph* to emphasize the digraph representation of a coevolutionary problem. It should be clear from the context of subsequent presentations whether we are referring to coevolutionary problems or just the digraphs as mathematical structures.

For this study, we consider two general classes of coevolutionary problems with their corresponding digraph representations.

**Definition 2.4.2.** Consider the coevolutionary problem where $V_{S(n)}$ is the solution set with $|V_{S(n)}| = n$. Its underlying graph $UG(D_C)$ is complete. Then: (i) a *coevolutionary tournament* is an orientation of $UG(D_C)$, and (ii) a *coevolutionary semicomplete digraph* is a biorientation of $UG(D_C)$. For each $n = 3, 4, 5, \ldots$, let $\mathcal{T}(V_{S(n)})$ and $\mathcal{SD}(V_{S(n)})$ be the sets of all tournaments and semicomplete digraphs of order $n$, respectively. $\mathcal{T}(V_{S(n)})$ and $\mathcal{SD}(V_{S(n)})$ represent the family of all coevolutionary tournaments and semicomplete digraphs with $n$ candidate solutions, respectively.[3] $\mathcal{SD}(V_{S(n)})$ generalizes $\mathcal{T}(V_{S(n)})$ since adjacent vertices in $V_{S(n)}$ can be completely bioriented. That is, $\mathcal{T}(V_S) \subset \mathcal{SD}(V_S)$.

---

[2]Cycles of lengths two and one can be defined as well and usually in context of digraphs that admit specific connectivity structures. A 2-cycle is the case whereby two adjacent vertices $u, v \in V$ is completely bioriented via $uv$ and $vu$. A 1-cycle is simply the case of a self-loop $uu$.

[3]Implicit in our notation are the corresponding sets of arcs in $\mathcal{T}(V_{S(n)})$ and $\mathcal{SD}(V_{S(n)})$ satisfying constraints for tournaments and semicomplete digraphs, respectively.

To provide a direct context, there is a wide range of coevolutionary problems that can be modelled as *two-player strategic games*. In each case, the game consists of a set of *pure strategies* and a *dominance relation* on those strategies. Both players have access to the same *strategy set*. Each player will select a single pure strategy from the strategy set and then implements that strategy during game-play. The game outcome is decided by the dominance relation on the two pure strategies implemented for the game-play. Some games have only *win-lose* game outcomes – the winning strategy is said to dominate the losing strategy. Other games (two-player board games such as chess) have a *win, lose, or draw* outcome. Digraphs naturally represent these coevolutionary problems. The set of vertices of the digraph is the strategy set. The dominance relation is captured and described by the set of arcs. We do not consider game-plays between copies of the same pure strategy for simplicity – there is no self-loop in the digraph.

**Definition 2.4.3.** For any pair of strategies $u, v \in V_S$, the *winning strategy $v$ dominates* a *losing strategy $u$*. This relation is represented by the orientation $v \leftarrow u$ (or $u \rightarrow v$ if we say $u$ is *dominated* by $v$).[4] A draw is represented by the complete biorientation $u \rightleftarrows v$.

The notion of domination can be extended to subsets of $V_S$ (a vertex is a subset of $V_S$ with a single element). Consider two disjoint subsets of $V_S$, $V_S^1$ and $V_S^2$, whereby $V_S^1 \cap V_S^2 = \varnothing$. $V_S^2 \Leftarrow V_S^1$ denotes that all vertices $v \in V_S^2$ dominate all vertices $u \in V_S^1$. We say $V_S^2$ is the *dominant* subset with respect to $V_S^1$. Conversely, $V_S^1 \Rightarrow V_S^2 \equiv \forall v \in V_S^1 \ \forall u \in V_S^2 \ (v \rightarrow u)$. The tournament provides a representation for *all* strategic games with win-lose game outcomes. Semicomplete digraph as a further generalization of the tournament can be used to represent any strategic game with win, lose, or draw outcomes.[5]

## 2.5  Structures and Higher Organizations in Coevolutionary Digraphs

In practice, one often is confronted with a very large coevolutionary digraph with little (connectivity) information. A coevolutionary tournament with $|V|$ candidate solutions has $2^{(|V|)(|V|-1)/2}$ pairs of comparisons. This exponential increase in pairwise comparisons is further compounded by a high (and possibly exponential) growth rate in the number of *parametric* representations of candidate solutions. Consider a simple *normal form* game such as the *iterated prisoner's dilemma* restricted to just memory-one, reactive (self and opponent's past plays) strategies. If one uses a direct look-up table representation, a $q$-choice game would generate $q^{q^2+1}$ different combinations of strategy tables (Chong et al., 2012). A two-choice game would generate eight strategy tables and 784 pairwise comparisons, which increase to at least $6.18 \times 10^{28}$ strategy tables and $3.66 \times 10^{114}$ pairwise comparisons for a six-choice game in a full tournament.

Performing a brute-force computation on a very large coevolutionary digraph (e.g. decide if the problem has a single dominant solution) quickly becomes a very expensive task. Instead, we study specific structures leading to some form of higher organizations in coevolutionary digraphs that can provide useful characterizations of coevolutionary problems. These structures emerge from constraints that could be placed on the biorientations of coevolutionary digraphs.

---

[4]Standard digraph terminology define domination in the reverse order, i.e., $u$ *dominates* $v$ is denoted as $u \rightarrow v$. Our choice to define domination in this manner becomes clear in the next section when we discuss coevolutionary processes on digraphs and their Markov chain models.

[5]We are not describing how one can find the best strategy but how relations of strategies can be captured, e.g., *transitive* strategic games represented by *acyclic orientations* of complete graphs.

## 2.6 Connectivity Structures in Digraphs

We first define basic connectivity structures in digraphs before introducing qualitative characterizations of tournaments having specific connectivity conditions.

**Definition 2.6.1.** A digraph $D$ is said to be *strongly connected* (or *strong*) if for each pair of distinct vertices $u, v \in V(D)$ there exist both $(u, v)$-path and $(v, u)$-path.

Given a pair of distinct vertices $u, v \in V(D)$, $v$ is said to be *reachable* from $u$ if and only if there is an $(u, v)$-path in $D$. A vertex is reachable from itself and so by definition a digraph with a single vertex is strong. A strong digraph will have each vertex reachable from all other vertices. There exists a closed Hamilton walk that cover all vertices in the digraph. The following describes strong digraphs having specific cycle structures.

**Definition 2.6.2.** A digraph $D = (V, A)$ of order $n$ that has $k$-cycles of all lengths $k = 3, 4, 5, \ldots, n$ is said to be *pancyclic*. The digraph is further said to be *vertex-pancyclic* if for every $v \in V(D)$ and every $k \in \{3, 4, 5, \ldots, n\}$, there is a $k$-cycle containing $v$.

*Acyclic* digraphs are those without cycles and are defined formally as follows.

**Definition 2.6.3.** Let $v_1, v_2, v_3, \ldots, v_n$ be an ordering of $V(D)$. If $v_1, v_2, v_3, \ldots, v_n$ is such that for all arcs $(v_i, v_j) \in A(D)$, $i < j$, the digraph $D = (V, A)$ is said to have an *acyclic ordering*. Every acyclic digraph has a vertex with zero in-degree $d_D^-(v_1) = 0$ and a vertex of zero out-degree $d_D^+(v_n) = 0$.

Every acyclic digraph has an acyclic ordering (Bang-Jensen and Gutin, 2009).

**Definition 2.6.4.** A digraph $D = (V, A)$ is *transitive* if for every pair of distinct arcs $(u, v), (v, w) \in A(D)$, $u \neq w$ and $(u, w) \in A(D)$.

A tournament is transitive if and only if it is acyclic (Bang-Jensen and Gutin, 2009).

For each $n = 3, 4, 5, \ldots$, let $\mathcal{T}(V_{S(n)})$ be the set of all tournaments of order $n$. Vertices for each tournament $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ can be labelled (indexed) in a manner so that $v_1, v_2, v_3, \ldots, v_n$ is ordered according to $d_T^-(v_1) \leq d_T^-(v_2) \leq d_T^-(v_3) \leq \cdots \leq d_T^-(v_n)$. This is done using the score sequence.

**Definition 2.6.5.** The *score sequence* of a tournament $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ is the $n$-tuple $(s_1, s_2, s_3, \ldots, s_n)$ with the constraints $s_1 \leq s_2 \leq s_3 \leq \cdots \leq s_n$. Each $s_i$, $i = 1, 2, 3, \ldots, n$, corresponds to the number of vertices that $v_i$ dominates, i.e., $s_i = d_T^-(v_i)$.

The following two definitions (Moon, 1968) are crucial in the next subsection where we present our results on how vertices in a digraph can be grouped or classified in some manner that reveal useful structures.

**Definition 2.6.6.** Consider a *vertex partition* of vertices $V_{S(n)}$ of a tournament $T(V_{S(n)}, A)$ into two disjoint, nonempty subsets $V_{S(n)}^1$ and $V_{S(n)}^2$ whereby $V_{S(n)}^1 \cap V_{S(n)}^2 = \varnothing$ and $V_{S(n)}^1 \cup V_{S(n)}^2 = V_{S(n)}$. $T(V_{S(n)}, A)$ is *reducible* if it can be vertex-partitioned such that $V_{S(n)}^1 \Rightarrow V_{S(n)}^2$ ($V_{S(n)}^2$ is the dominant subset). Otherwise, it is *irreducible*.

Some reducible digraphs further admit vertex partition (decomposition) into multiple, disjoint subdigraphs, which is defined as follows.

**Definition 2.6.7.** A digraph $D$ can be *decomposed* into subdigraphs $D^{(1)}, D^{(2)}, D^{(3)}, \ldots, D^{(l)}$ where $V(D^{(1)}) \cup V(D^{(2)}) \cup V(D^{(3)}) \cup \cdots \cup V(D^{(l)}) = V(D)$ and $V(D^{(i)}) \cap V(D^{(j)}) = \varnothing$ when $i \neq j$.

## 2.7 Qualitative Characterizations of Coevolutionary Digraphs

We first summarize several classic results of tournament theory on connectivity structures in tournaments. The results will be presented in the original form of the (referenced) primary texts (slightly restated for consistency with our notations).

**Theorem 2.7.1** (Rado (1943) (Moon, 1968))**.** *A tournament $T(V_{S(n)}, A)$ is strong if and only if it is irreducible.*

**Theorem 2.7.2** (Camion (1959) (Bang-Jensen and Gutin, 2009))**.** *Every strong tournament is hamiltonian.*

**Theorem 2.7.3** (Moon (1968) (Moon, 1968))**.** *Let $T(V_{S(n)}, A)$ be a strong tournament on $n \geq 3$ vertices. For every $v \in V_{S(n)}(T)$ and every integer $k \in \{3, 4, 5, \ldots, n\}$, there exists a $k$-cycle through $v$ in $T(V_{S(n)}, A)$. In particular, $T(V_{S(n)}, A)$ is hamiltonian if and only if it is strong.*

We exploit these theoretical results and provide useful qualitative characterizations to the family of coevolutionary tournaments $\mathcal{T}(V_{S(n)})$. In this paper, we state and show that a coevolutionary problem is either solvable (its corresponding tournament is reducible) or not (the tournament is irreducible).

**Lemma 2.7.4.** *A coevolutionary tournament $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ on $n \geq 3$ is either reducible or irreducible.*

*Proof.* This is simply a direct consequence of the definition of reducibility and irreducibility of tournaments (as given by Definition 2.6.6).[6]  $\square$

We next show that a solvable coevolutionary problem has a subset of solutions that are preferred over the remaining solutions.

**Lemma 2.7.5.** *A reducible coevolutionary tournament $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ of order $n \geq 2$ can be vertex-partitioned into two disjoint nonempty subsets $V^1_{S(n)}$ and $V^2_{S(n)}$. The dominant subset $V^2_{S(n)}$ contains solutions that dominate all remaining solutions in the other subset $V^1_{S(n)} = V_{S(n)} - V^2_{S(n)}$. The minimal dominant subset contains a single solution.*

*Proof.* The first part of the lemma is simply a direct consequence of the definition of reducibility of tournaments. It remains to prove the second part. We first make note that every tournament can be decomposed into subtournaments $T^i (i = 1, 2, 3, \ldots, l)$ whereby (Moon, 1968):

a) Every vertex in $T^{(j)}$ dominates every vertex in $T^{(i)}$ if $1 \leq i < j \leq l$.

b) Every subtournament $T^{(i)}$ is either irreducible or transitive.

c) No two consecutive subtournaments $T^{(i)}$ and $T^{(i+1)}$ are both transitive.

Consider the trivial case whereby the tournament is transitive, $l = 1$. As a consequence of Theorem 2.7.3, a tournament $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ of order $n > 2$ is transitive if and only if it is acyclic. Obviously, we can always vertex partition a transitive tournament by taking the vertex with in-degree $n-1$ as $V^2_{S(n)}$ and all the remaining $n-1$ vertices as $V^1_{S(n)}$ so that $V^1_{S(n)} \Rightarrow V^2_{S(n)}$. For a reducible tournament, assume that it can be decomposed into two subtournaments ($l = 2$). Since the two subtournaments $T^{(1)}$ and $T^{(2)}$ cannot be both transitive, then one of them must be irreducible. Let $V(T^{(1)}) \Rightarrow V(T^{(2)})$. $V(T^{(2)})$ with a single vertex is by definition strong and as such irreducible.  $\square$

---

[6] A coevolutionary tournament with one vertex is irreducible since it is a strong tournament by definition. A tournament with two vertices is reducible since there is always one vertex that dominate the remaining vertex in the tournament.

We only need to show the case above where the dominant subset is minimal. There are other cases as well, for example, $T^{(2)}$ is transitive and $T^{(1)}$ is irreducible. Here, even though $V(T^{(1)}) \Rightarrow V(T^{(2)})$, the single dominant vertex in a transitive $T^{(2)}$ obviously dominates all other vertices in $T^{(2)}$ and $T^{(1)}$ as well. In another case, we have both irreducible subtournaments. For these coevolutionary tournaments, the maximal dominant subset that induces $T^{(2)}$ and for which $V(T^{(1)}) \Rightarrow V(T^{(2)})$ is of order $n-1$.

For coevolutionary problems that are not solvable, we have the following result.

**Lemma 2.7.6.** *An irreducible coevolutionary tournament $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ is strongly connected. It is hamiltonian and furthermore, vertex-pancyclic.*

*Proof.* This is simply a direct consequence of Theorems 2.7.1, 2.7.2, and 2.7.3. □

The main implication is that an irreducible coevolutionary tournament does not have a subset of dominant solutions. In addition to every $u, v, w \in V_{S(n)}(T)$ contained in a 3-cycle, this coevolutionary tournament is vertex-pancyclic. One can find the solution $v_n$ that dominates all other solutions $v_2, v_3, v_4, \ldots, v_{n-1}$ but is dominated by $v_1$, which in turn is dominated by $v_2, v_3, v_4, \ldots, v_{n-1}$ (see Figure 1). This is a consequence of the Hamilton cycle in the tournament. In the context of a game and optimality of a strategy that is defined in terms of the number of opponent strategies it defeats, one finds the situation whereby the *best strategy defeats all but the worst strategy*. The lack of dominant solutions becomes more apparent in the case of *regular* coevolutionary tournaments whereby each solution $v$ has the same score with $d_D^-(v) = d_D^+(v)$, which generalize $T(V_{S(3)}, A)$ with $u \to v \to w \to u$ (as in the game *rock-paper-scissors*).
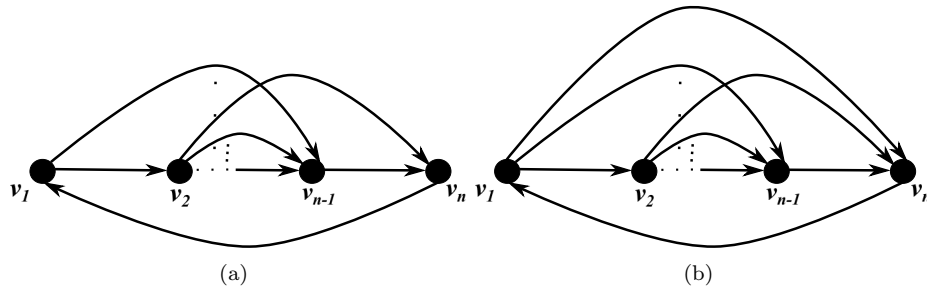


Figure 1: A vertex-pancyclic coevolutionary (a) tournament and (b) semicomplete digraph of order $n$.

Our results can be extended to the family of coevolutionary semicomplete digraphs $\mathcal{SD}(V_S)$ as well. The arguments behind Theorems 2.7.1, 2.7.2, and 2.7.3 for tournaments can also be made for their simple generalization in the form of semicomplete digraphs. For example, Theorem 2.7.3 on strong connectivity and vertex pancyclicity for tournaments has been restated for semicomplete digraphs in (Bang-Jensen et al., 1997). The original argument (Moon, 1968) that makes use of a proof by induction of a $k$-cycle (where $3 \le k \le n$) containing a vertex $v$ and the existence of a $(k+1)$-cycle containing $v$ remains unchanged if one replaces instances of tournaments with semicomplete digraphs. One can make an observation (e.g., see Figures 1 (a) and (b)) that the introduction of 2-cycles does not make a semicomplete digraph vertex pancyclic. Instead, it arises from the presence of 3-cycles and the Hamilton cycle.

Unlike the case of irreducible coevolutionary digraphs, one can employ search procedures to discover the dominant subset in a reducible coevolutionary digraph. We first

introduce the following notion on strong connectivity of subdigraphs $D^{(i)}$ of $D$ (Bang-Jensen and Gutin, 2009). Then, we conclude with a useful description of structural properties in reducible coevolutionary digraphs that we apply later.

**Definition 2.7.7.** The *strong component* of $D$ is a maximal induced strong subdigraph of $D$. It is maximal in the sense that if a strong $D^{(i)}$ and strong $D^{(j)}$ are connected by arcs in both directions, in which case there exists both $(u,v)$-path and $(v,u)$-path for every $u \in D^{(i)}$ and $v \in D^{(j)}$, then they belong to the same strong subdigraph $D^{(k)} = D^{(i)} \cup D^{(j)}$. Obviously a single-vertex $D^{(i)}$ is strong by definition. A digraph $D$ can be decomposed into its strong components $D^{(1)}, D^{(2)}, D^{(3)}, \ldots, D^{(l)}$, with each strong component being distinct by definition of it being a maximal induced subdigraph.

**Lemma 2.7.8.** *A reducible coevolutionary semicomplete digraph* $SD(V_{S(n)}, A) \in \mathcal{SD}(V_{S(n)})$ *of order* $n \geq 2$ *has a strong decomposition* $V\big(D^{(1)}\big) \cup V\big(D^{(2)}\big) \cup V\big(D^{(3)}\big) \cup \ldots \cup V\big(D^{(l)}\big)$ *with a unique ordering* $D^{(1)}, D^{(2)}, D^{(3)}, \ldots, D^{(l)}$ *whereby* $D^{(i)} \Rightarrow D^{(j)}$ *when* $i < j$. $D^{(1)} \left( D^{(l)} \right)$ *is the initial (terminal) strong component. In particular, the dominant solutions of* $SD(V_{S(n)}, A)$ *are contained in its terminal strong component.*

*Proof.* As with the case of tournaments, we first note that a reducible semicomplete digraph can be decomposed into strong components $D^i (i = 1, 2, 3, \ldots, l)$. Let the strong components be ordered as $D^{(1)} \Rightarrow D^{(2)} \Rightarrow D^{(3)} \Rightarrow \ldots \Rightarrow D^{(l)}$ with $D^{(i)} \Rightarrow D^{(j)}$ when $i < j$ (Bang-Jensen et al., 1996). One can perform a *contraction* operation on the strong components that replaces $D^{(i)}$ with a single vertex $g\big(D^{(i)}\big)$ and remove all parallel arcs (Bang-Jensen and Gutin, 2009). When contraction is applied to all the strong components of $D$, the resulting digraph is called the *strong component digraph* $SC(D)$. Note that $SC(D)$ is now $g\big(D^{(1)}\big) \to g\big(D^{(2)}\big) \to g\big(D^{(3)}\big) \to \cdots \to g\big(D^{(l)}\big)$ with $g\big(D^{(i)}\big) \to g\big(D^{(j)}\big)$ when $i < j$, i.e., it has an acyclic ordering. The *initial (terminal)* strong component of $D$ is obviously the strong component whose corresponding vertex in $SC(D)$ has zero in-degree (out-degree). The same argument can be made for tournaments. However, with a reducible tournament one may have to apply contraction on transitive subtournaments, in which case the resulting digraph will be acyclic as well. $\square$

## 2.8 Quantitative Characterizations of Coevolutionary Digraphs

In demonstrating several structural results relating to cycle structures in coevolutionary digraphs crucial to characterizing their dichotomy in solvability, we make no assumption on availability of prior information relevant to uncovering those structures. Where more information becomes available, we can make quantitative characterizations relevant to the *complexity* of those cycle structures and further uncover their higher organizations. In particular, we first study how those cycles of various lengths can be *packed* in a manner that results in a strong decomposition of the coevolutionary tournament. We present the following result.

**Lemma 2.8.1.** *Let* $(s_1, s_2, s_3, \ldots, s_n)$ *be the sequence of integers satisfying* $s_1 \leq s_2 \leq s_3 \leq \cdots \leq s_n$. $(s_1, s_2, s_3, \ldots, s_n)$ *is the score sequence of a coevolutionary tournament* $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ *of order* $n$ *if and only if*

$$\sum_{i=1}^{k} s_i \geq \binom{k}{2}, \ k = 1, 2, 3, \ldots, n, \ \text{with equality when } k = n.$$

*In particular,* $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ *of order* $n \geq 3$ *is irreducible if and only if for each*

$k$, $1 \le k \le n - 1$

$$\sum_{i=1}^{k} s_i > \binom{k}{2} \ and \ \sum_{i=1}^{n} s_i = \binom{n}{2}.$$

*Otherwise, $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ is reducible whereby for some $k$, $1 \le k \le n - 1$*

$$\sum_{i=1}^{k} s_i = \binom{k}{2}.$$

*Proof.* The first statement on score sequences of coevolutionary tournaments is a direct consequence of a theorem by Landau (1953) (Bang-Jensen and Gutin, 2009).[7] The second statement on score sequences of irreducible coevolutionary tournament is a direct consequence of a theorem by Moser and Harary (1966) (Bang-Jensen and Gutin, 2009). The final statement immediately follows from the second statement, thereby providing conditions for reducibility of coevolutionary tournaments (Moon, 1963). □

Lemma 2.8.1 makes the only assumption of a labelled tournament (its score sequences are known). There are several consequences to this result. A coevolutionary tournament of order $n$ is transitive if it has a score sequence of $(0, 1, 2, \ldots, n-1)$ whereby $\sum_{i=1}^{k} s_i = \binom{k}{2}$ for each $k$, $1 \le k \le n$. Lemma 2.8.1 can be used to identify strong components in a reducible coevolutionary tournament. The idea is that to look for successive values of $k$ for which the equality $\sum_{i=1}^{k} s_i = \binom{k}{2}$ holds. This follows from Lemma 2.7.5 on the decomposition of tournaments into subtournaments $T^j (j = 1, 2, 3, \ldots, l)$. A strong subtournament $T^{(j)}$ will be sandwiched between strong or transitive subtournaments, $T^{(j-1)} \Rightarrow T^{(j)} \Rightarrow T^{(j+1)}$.

One can formalize this as follows. We consider a coevolutionary tournament $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ with its solution set $V_{S(n)}$ indexed according to its score sequence. Then, let $\sum_{i=1}^{p} s_i = \binom{p}{2}$, $\sum_{i=1}^{q} s_i = \binom{q}{2}$, and $\sum_{i=1}^{k} s_i > \binom{k}{2}$ for $p + 1 \le k \le q - 1$ with $0 \le p < q \le n$. The subtournament induced by the subset $\{v_{p+1}, v_{p+2}, v_{p+3}, \ldots, v_q\}$ is a strong subtournament $T^{(j)}$. Furthermore, the score sequence that is restricted to the strong subtournament $T^{(j)}$ itself is given by $(s_{p+1} - p, s_{p+2} - p, s_{p+3} - p, \ldots, s_q - p)$ (Pirzada, 2012). If $T(V_{S(n)}, A)$ is irreducible, the procedure will just return a single strong component – the entire tournament itself.

Finally, we will exploit Lemma 2.8.1 to construct an index relevant to characterizing complexity of cycle structures in coevolutionary digraphs. We give a precise definition for the index as follows.

**Definition 2.8.2.** Consider a coevolutionary tournament $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ of order $n$. The vertices of $T(V_{S(n)}, A)$ is indexed as $v_1, v_2, v_3, \ldots, v_n$ according to the score sequence $(s_1, s_2, s_3, \ldots, s_n)$. Let the two integers $i, j = 1, 2, 3, \ldots, n$. The tournament is associated with two sequences $\mathcal{S} = (s_1, \sum_{i=1}^{2} s_i, \sum_{i=1}^{3} s_i, \ldots, \sum_{i=1}^{n} s_i)$ and $\mathcal{L} = \left(0, \binom{2}{2}, \binom{3}{2}, \ldots, \binom{n}{2}\right)$. We write the $i$th element of the $\mathcal{S}$ as $\mathcal{S}_i$ (similarly, for $\mathcal{L}$ we write $\mathcal{L}_i$). Then, we define the index $\nu$ as follows

$$\nu = \sum_{i=1}^{n} \mathbb{I}\{\mathcal{S}_i = \mathcal{L}_i\}, \tag{1}$$

---

[7]Landau's theorem can be generalized further. For example, Reid and Zhang provided in (Reid and Zhang, 1998) necessary and sufficient conditions for a sequence of integers to be the score sequence of semicomplete digraphs. Moon (1963) showed that the original theorem can be extended to the case whereby one can assign a real number $[0, 1]$ to the outcome of the pairwise comparison (game) between a pair of vertices (strategies). That is, for a game between two strategies $i$ and $j$, $\alpha_{ij}$ is the score for $i$ and $\alpha_{ji} = 1 - \alpha_{ij}$ is the score for $j$.

where

$$\mathbb{I}\{\mathcal{S}_i = \mathcal{L}_i\} = \begin{cases} 1 & \text{if } \sum_{j=1}^{i} s_j = \binom{i}{2} \\ 0 & \text{if } \sum_{j=1}^{i} s_j \neq \binom{i}{2}. \end{cases}$$

The index $\nu$ ranges from 1 to $n$. Essentially, it is a count of the number of strong components in the coevolutionary tournament. When this index is used together with structural results such as Lemma 2.7.8 that provide qualitative descriptions of tournament decomposition, it provides a useful macro-scale indication of cycle structure complexity in coevolutionary tournaments that is summarized as a simple quantity. $\nu = 1$ if and only if the tournament is strong, while $\nu = n$ if and only if the tournament is transitive (each vertex is a strong component and all these vertices have an acyclic ordering). A value of $\nu$ closer to $n$ indicates the presence of many shorter length cycles that can be grouped separately into different subtournaments. When $\nu$ is close to one, the tournament has longer length cycles between $n/2$ to $n-1$. $\nu$ can be computed quickly (the computation cost is linear with the order of the tournament with precomputed $\mathcal{L}$s). Later, we demonstrate how this index is used in computational studies.

## 3 Coevolutionary Processes on Digraphs

We introduce the second part of our framework and show that coevolutionary processes can be formulated naturally as sampling processes on digraphs. For simplicity and to allow for in-depth analysis, we consider population-one coevolutionary systems, which corresponds to random walks on digraphs. These would cover various coevolutionary systems such as those that are used in the context of learning game-playing strategies through self play. We will formulate these random walks as discrete time Markov chains operating on coevolutionary digraphs. This will allow us to study how specific structures in coevolutionary digraphs can affect the dynamics of coevolutionary processes.

### 3.1 Coevolutionary Processes as Random Walks on Digraphs

Algorithm 1 describes a standard random walk on labelled digraphs $D$. It starts at a random vertex and in each subsequent step, randomly moves to an out-neighbour of the current vertex. This requires prior information on connectivity structures (out-neighbourhoods of all vertices in $D$). In practice, one would implement a random walk on unlabelled digraphs such as the (1+1)CEA (Algorithm 2). In the (1+1)CEA, the random move in each step is made through a two-stage process of variation (generates a random *offspring* from the neighbourhood of the *parent* vertex) and selection (moves to the offspring vertex if it dominates the parent vertex). Unlike a random walk, there is a chance the (1+1)CEA stays at the current vertex in a step because it chooses one of its in-neighbours (the probability is proportionate to the order of the in-neighbourhood). Note as well that in practice a (1+1)CEA can apply an arbitrary stopping criterion (finite number of steps as in Algorithm 2).

### 3.2 Formulating the Coevolutionary Markov Chain

Regardless of how random walks on digraphs are implemented algorithmically, our focus is on the underlying *stochastic processes* that generate those walks. We can exploit various tools for quantitative analysis of coevolutionary processes consistent with the qualitative characterizations we have presented earlier on coevolutionary problems. We first motivate formulating coevolutionary processes on digraphs as a specific class of stochastic processes – discrete time Markov chains. We will refer readers to (Norris, 1998)

---

**Algorithm 1** Random Walk on Digraph

---

1: **procedure** RWALKD($D, u$)
2:     **Initialization**
3:     $v \in V(D)$ chosen (uniformly) at random.
4:     $u := v$.
5:     **repeat**
6:         **if** $N_D^+(u) \neq \varnothing$ **then**
7:             $v \in N_D^+(u)$ chosen uniformly at random.    $\triangleright$ With probability $1/|N_D^+(u)|$.
8:         **end if**                         $\triangleright$ Else $u$ remains as it is.
9:         $u := v$.
10:     **until** Forever
11: **end procedure**

---

**Algorithm 2** (1+1)CEA on Digraph $D$

---

1: **procedure** CEA($D, u, t_{\text{stop}}$)
2:     **Initialization**
3:     $t := 0$.
4:     $v_t \in V(D)$ chosen (uniformly) at random.
5:     $u := v_t$.
6:     **Generational Process**
7:     **while** $t \leq t_{\text{stop}}$ **do**
8:         **Variation**
9:         $v_t \in V(D) - u$ chosen uniformly at random.
10:         **Selection**
11:         **if** $v_t \leftarrow u$ **then**
12:             $u := v_t$.
13:         **end if**                    $\triangleright$ Else $u$ remains as it is.
14:         $t := t + 1$.
15:     **end while**
16: **end procedure**

---

as an introductory text and (Meyn and Tweedie, 2009) for more advance materials in Markov chains. We will provide relevant materials to ensure the paper is self-contained.

**Definition 3.2.1.** A *discrete time stochastic process* $\boldsymbol{\Phi}$ is a collection of $\{\Phi_t : t \in \mathbb{N}_0\}$ of random variables with each $\Phi_t$ taking values in a countable set $X$ at time $t$. $\mathbb{N}_0$ is the set of natural numbers that include 0. A *random walk* is a discrete time process $\boldsymbol{\Phi} = \{\Phi_t : t \in \mathbb{N}_0\}$ that takes a step on the digraph from $\Phi_t$ to $\Phi_{t+1}$. Each $\Phi_t$ takes values from the *state space* $X$ corresponding to the finite set of vertices $V(D)$ of the digraph $D = (V, A)$. Each state $x \in X$ is the distinct vertex $v \in V$.

Constructing coevolution as a discrete time process is technically involved as one needs to work on probability spaces. In general, one first prescribes the structural definition for the sample path behaviors and the dynamics of steps on the digraph through time to define the probability laws that govern the evolution of the chain $\boldsymbol{\Phi}$. Even in the usual case where sequences $\{\Phi_0, \Phi_1, \Phi_2, \ldots\}$ take *discrete* values in $X$, developing a complete, functional description of the distribution of the random walk on most digraphs other than those with structures that strictly constrain its sample path behaviors (e.g., acyclic tournaments) is difficult. Instead, we use the observation that coevolutionary

processes satisfy the *Markov property* – the step to a future state only depends on the current state.[8] We exploit this to define and construct the probabilistic motion of the coevolutionary process on digraphs.

We first note the following definition and facts of a Markov chain.

**Definition 3.2.2.** The stochastic process $\boldsymbol{\Phi} = \{\Phi_t : t \in \mathbb{N}_0\}$ is a Markov chain if for every $n$ and any sequence $\{\Phi_0, \Phi_1, \Phi_2, \ldots, \Phi_n\}$,

$$\mathbb{P}_{\boldsymbol{\mu}}(\Phi_{n+1} = x_{n+1} \mid \Phi_n = x_n, \Phi_{n-1} = x_{n-1}, \ldots, \Phi_0 = x_0) = \mathbb{P}(x_n, x_{n+1}), \qquad (2)$$

given $\boldsymbol{\mu}$ is the initial distribution of the Markov chain.[9]

One can immediately observe that this definition incorporates two main characteristics: (i) *time homogeneity* (Theorem 1.1.1, page 2 in (Norris, 1998)) whereby

$$\mathbb{P}_{\boldsymbol{\mu}}(\Phi_0 = x_0, \Phi_1 = x_1, \Phi_2 = x_2, \ldots, \Phi_n = x_n) = \mu(x_0)\mathbb{P}(x_0, x_1)\mathbb{P}(x_1, x_2)\cdots\mathbb{P}(x_{n-1}, x_n), \quad (3)$$

with the probabilities $\mathbb{P}(x_t, x_{t+1})$ depending only on the pairs $(x_t, x_{t+1})$, and (ii) *memory loss* (Theorem 1.1.2, page 3 in (Norris, 1998)) whereby the probabilities $\mathbb{P}(x_t, x_{t+1})$ are independent of the time step $t$.

We formulate and define the Markov chain for a coevolutionary process on digraphs as follows.

**Definition 3.2.3.** Let $D(V_{S(n)}, A) \in \mathcal{D}(V_{S(n)})$ be a coevolutionary digraph having a finite vertex set $V_{S(n)}$. A coevolutionary Markov chain is a random walk on $D(V_{S(n)}, A)$ with initial distribution $\boldsymbol{\mu}$ over and Markov transition matrix $\boldsymbol{P}$ on the state space $X = V_{S(n)}$ satisfying

(i) $\boldsymbol{\mu} = (\mu_x : x \in X)$, where $0 \le \mu_x \le 1$ and $\sum_{x \in X} \mu_x = 1$, and

(ii) $\boldsymbol{P} = (\mathbb{P}(x, z) : x, z \in X)$, where for every row is a distribution whereby $0 \le \mathbb{P}(x, z) \le 1$ and $\sum_{y \in X} \mathbb{P}(x, y) = 1$.

Guarantees that such a *coevolutionary Markov chain* $\boldsymbol{\Phi} = \{\Phi_t : t \in \mathbb{N}_0\}$ defined on the state space $X = V_{S(n)}$, initial distribution $\boldsymbol{\mu}$ and transition matrix $\boldsymbol{P}$ are given by the *Existence Theorem* of Markov chains on countable state space (Chung, 1960). By definition a coevolutionary digraph $D(V_{S(n)}, A)$ has a discrete $V_{S(n)}$ with finite number of elements, which is a countable set.

The distribution describing a coevolutionary Markov chain is computed from its $\boldsymbol{\mu}$ and $\boldsymbol{P}$. We start the Markov chain with the initial distribution being fixed (point mass) at $x_0 \in X$. Subsequently, for each $x_0 \in X$, we define $\mathbb{P}_{x_0}$ of the Markov chain inductively:

$$
\begin{aligned}
\mathbb{P}_{x_0}(\Phi_0 = x_0) &= 1, \\
\mathbb{P}_{x_0}(\Phi_0 = x_0, \Phi_1 = x_1) &= \mathbb{P}(x_0, x_1), \\
\mathbb{P}_{x_0}(\Phi_0 = x_0, \Phi_1 = x_1, \Phi_2 = x_2) &= \mathbb{P}(x_0, x_1)\mathbb{P}(x_1, x_2),
\end{aligned}
$$

---

[8]Loosely speaking, the dynamics of the coevolutionary process as it walks on the digraph is such that it *restarts* in every step. For example, with the (1+1)CEA, the probability that it steps from a particular vertex $x_t = u \in V(D)$ to an out-neighbour $x_{t+1} = v \in N_D^+(u)$ is independent of the time (generational) step $t$, i.e., the probability is the same whether it is $x_0 = u$ at the start or $x_t = u$, which is $t$ generations later.

[9]Usually one will take the initial distribution $\boldsymbol{\mu}$ as the point mass $\delta_{x_0}$ at $x_0$. The point mass at $x$ is written as $\delta_x = (\delta_{xy} : y \in X)$ whereby

$$\delta_{xy} = \left\{ \begin{array}{ll} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y. \end{array} \right.$$

and so forth, using $\boldsymbol{P}$ that gives the chain's *one-step transition probabilities* $\mathbb{P}(x_t, x_{t+1})$ (Meyn and Tweedie, 2009). Other useful computations can be made, for example, the probabilities starting from $x$ after $n$ steps that the chain is at $y$, $\mathbb{P}_\mu(\Phi_n = y \mid \Phi_0 = x) = \mathbb{P}^n(x, y)$. One first sets $\boldsymbol{P}^0 = \mathbf{I}$ (the identity matrix), which corresponds to the initial conditions $\mathbb{P}^0(x, x) = 1$ for all $x \in X$. Then, the *n-step transition matrix* given by the iterates $\boldsymbol{P}^n = (\mathbb{P}^n(x, y) : x, y \in X)$ for $n \geq 0$ can be defined inductively through $\boldsymbol{P}^{n+1} = \boldsymbol{P}^n \boldsymbol{P}$ using the following expression

$$\mathbb{P}^{n+1}(x, z) = \sum_{y \in X} \mathbb{P}^n(x, y) \mathbb{P}(y, z). \tag{4}$$

## 3.3 Coevolutionary Markov Chain Examples

We construct four coevolutionary Markov chains (two for each case of transitive and regular tournaments). The initial distribution can be arbitrary so we just need to construct the appropriate Markov transition matrix. These examples highlight cases where we can do the constructions formally and later in Section 4.4 to perform quantitative analysis without resorting to sampling and large-scale computations as is often the case in real-world problems. The first example is the random walk on labelled transitive coevolutionary tournaments. Let $T(V_{S(m)}, A)$ be a transitive coevolutionary tournament with vertices $(v_1, v_2, \ldots, v_m)$ indexed according to its score sequence $(s_1, s_2, \ldots, s_m)$. We simplify our notation and let $\boldsymbol{P}_{m \times m} = (p_{ij} : i, j \in \{1, 2, \ldots, m\})$ be the Markov transition matrix for this coevolutionary Markov chain on $T(V_{S(m)}, A)$.
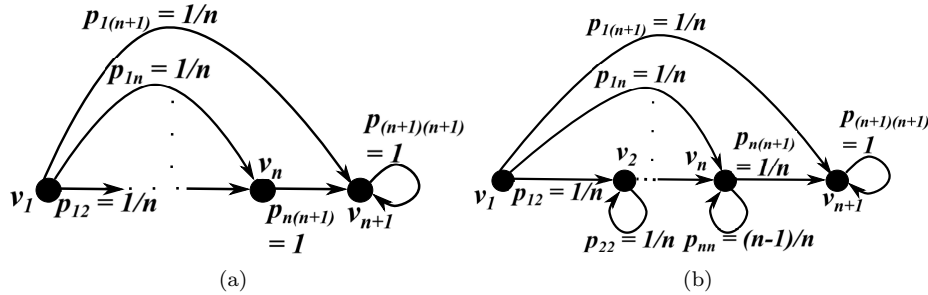


Figure 2: The coevolutionary Markov chain for a (a) random walk on labelled transitive tournaments $T(V_{S(n+1)}, A)$ and (b) (1+1)CEA on transitive tournaments $T(V_{S(n+1)}, A)$.

$\boldsymbol{P}_{m \times m}$ is defined inductively as follows. The basis $(n = 0)$ is $T(V_{S(m)}, A)$ of order $m = n + 1 = 1$. Since $T(V_{S(m)}, A)$ has one vertex that is also its winning vertex, $v_m = v_1$, it has a corresponding transition probability corresponding to the self-loop $p_{mm} = 1$. As such, $\boldsymbol{P}_{1 \times 1} = [p_{11}] = [1]$. The induction step $(n + 1)$ adds an additional vertex so that the tournament is of order $m = n + 1$. This additional vertex is indexed $v_1$. There are $n = m - 1$ vertices indexed from $v_2$ to $v_m$ that dominates $v_1$ (see Figure 2(a)). This gives equal probability of transitions from $v_1$ to each of $\{v_2, v_3, \ldots, v_m\}$ so that

$$
\begin{aligned}
p_{11} &= 0, \\
p_{1j} &= 1/(m-1), \quad j = 2, 3, \ldots, m, \\
p_{i1} &= 0, \qquad\qquad i = 2, 3, \ldots, m.
\end{aligned}
$$

We can then construct the Markov transition matrix for the random walk on labelled

transitive tournaments $T(V_{S(n+1)}, A)$ of order $n+1$ inductively on $m = 1, 2, \ldots, n+1$

$$
\begin{array}{rcll}
p_{ii} & = & 0, & i = m \\
p_{ij} & = & 1/(n+1-m), & i = m, \ j = i+1, i+2, \ldots, n+1, \\
p_{ij} & = & 0, & i = j+1, j+2, \ldots, n+1, \ j = m,
\end{array}
$$

and $p_{(n+1)(n+1)} = 1$. This will result in

$$
\boldsymbol{P}_{(n+1)\times(n+1)} = \left(
\begin{array}{ccccccc}
0 & \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} & \frac{1}{n} & \frac{1}{n} \\
0 & 0 & \frac{1}{n-1} & \cdots & \frac{1}{n-1} & \frac{1}{n-1} & \frac{1}{n-1} \\
0 & 0 & \ddots & \ddots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \frac{1}{3} & \frac{1}{3} \\
0 & 0 & 0 & \ldots & 0 & \frac{1}{2} & \frac{1}{2} \\
0 & 0 & 0 & \ldots & 0 & 0 & 1 \\
\hline
0 & 0 & 0 & \ldots & 0 & 0 & 1
\end{array}
\right),
\tag{5}
$$

where the last row indicates transition probabilities for the tournament's single dominant vertex, with a drawn horizontal line for emphasis.

The next example is the (1+1)CEA on unlabelled transitive coevolutionary tournament. Previously, the Markov transition matrix has a nice inductive structure (the lower right submatrix induces a random walk on the corresponding labelled subtournament). Here, the (1+1)CEA requires that one computes its Markov transition matrix for transitive coevolutionary tournaments of different orders. Fortunately, the computation for the transition probabilities is straightforward. Consider a transitive coevolutionary tournament $T(V_{S(n+1)}, A)$ of order $n+1$. The transition probabilities are given as follows:

$$
\begin{array}{lrcll}
\text{Upper triangle:} & p_{ij} & = & 1/n, & i < j+1, \ i = 2, 3, \ldots, n+1, \\
\text{Diagonal:} & p_{ii} & = & m/n, & m = i-1, \ i = 2, 3, \ldots, n+1, \\
\text{Lower Triangle:} & p_{ij} & = & 0, & i > j-1, \ i = 2, 3, \ldots, n+1,
\end{array}
$$

$$
\boldsymbol{P}_{(n+1)\times(n+1)} = \left(
\begin{array}{ccccccc}
0 & \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} & \frac{1}{n} & \frac{1}{n} \\
0 & \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} & \frac{1}{n} & \frac{1}{n} \\
0 & 0 & \frac{2}{n} & \ddots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \frac{1}{n} & \frac{1}{n} & \frac{1}{n} \\
0 & 0 & 0 & \ldots & \frac{n-2}{n} & \frac{1}{n} & \frac{1}{n} \\
0 & 0 & 0 & \ldots & 0 & \frac{n-1}{n} & \frac{1}{n} \\
\hline
0 & 0 & 0 & \ldots & 0 & 0 & 1
\end{array}
\right),
\tag{6}
$$

with the corresponding Markov chain illustrated in Figure 2(b).

We now consider two irreducible coevolutionary Markov chains. Let $T(V_{S(n)}, A)$ be a regular coevolutionary tournament of odd order $n \geq (3 + 2m)$, $m = 0, 1, 2, \ldots$. Its *adjacency matrix* $\mathbf{M}_{\mathrm{adj}}$ can be constructed as follows (Brualdi, 2010):

$$
\mathbf{M}_{\mathrm{adj}} = \mathbf{L}_n + \mathbf{L}_n^2 + \mathbf{L}_n^3 + \ldots + \mathbf{L}_n^{(n-1)/2},
$$

where $\mathbf{L}_n$ is the permutation matrix that is obtained from the permutation vector $(2, 3, 4, \ldots, n, 1)$. The permutation vector $(r_1, r_2, r_3, \ldots, r_m)$ specifies the row permutation operations on a $m \times m$ identity matrix. For the first row, 1 in the first column is

exchanged with 0 in the $r_1$-th column, and so forth, ending with the $m$-th row where 1 in the $m$-th column is exchanged with 0 in the $r_m$-th column.[10]

The value $\text{adj}(i, j) = 1$ in the adjacency matrix specifies the arc $i \rightarrow j$ of the corresponding digraph. The Markov transition matrix requires that we calculate from the digraph's *domination matrix* $\mathbf{M}_{\text{dom}}$ where $\text{dom}(i, j) = 1$ means $i \leftarrow j$. Note that $\mathbf{M}_{\text{dom}}$ is the transpose of $\mathbf{M}_{\text{adj}}$, $\mathbf{M}_{\text{dom}} = \mathbf{M}_{\text{adj}}^T$. We can then compute the Markov transition matrix for a random walk on labelled regular coevolutionary tournament as

$$\boldsymbol{P} = \frac{2}{(n-1)} \mathbf{M}_{\text{dom}}, \tag{7}$$

while for $(1 + 1)$CEA on unlabelled regular coevolutionary tournament, it is

$$\boldsymbol{P} = \frac{1}{2} \mathbf{I} + \frac{1}{(n-1)} \mathbf{M}_{\text{dom}}. \tag{8}$$

## 4 Quantitative Characterizations of Coevolutionary Markov Chains

In this section, we first demonstrate the link between structural properties in the coevolutionary digraphs and their impact on coevolutionary processes that operate on those digraphs. Subsequently, we will introduce and develop quantitative characterizations that are relevant to those coevolutionary Markov chains.

### 4.1 Class Structures

One fundamental notion that can help us understand the dynamics of a coevolutionary Markov chain is the *communication* classes on its state space. We can analyze a large and complex coevolutionary Markov chain by breaking it into smaller pieces having shared properties (states within them are equivalent in some sense). Let $D(V_{S(n)}, A)$ be a coevolutionary digraph. $\boldsymbol{P}$ is the Markov transition matrix of the coevolutionary Markov chain operating on $D(V_{S(n)}, A)$. Let $x$ and $y$ be two distinct states $x, y \in X = V_{S(n)}$.

**Definition 4.1.1.** $x \rightsquigarrow y$ ($x$ *leads to* $y$) if and only if there is a positive probability that starting at $x$ the Markov chain reaches $y$ after $n$-steps, $p_{xy}^{(n)} > 0$ for some $n \geq 0$.

The relation $\rightsquigarrow$ is transitive given the property of the $n$-step transition probability (as indicated by Equation 4) so that one can compute $p_{xz}^{(m+n)} \geq p_{xy}^{(m)} p_{yz}^{(n)} > 0$ for some $m \geq 0$ and $n \geq 0$. By definition, $x \rightsquigarrow x$, in which case the relation is reflexive.

**Definition 4.1.2.** $x \leftrightsquigarrow y$ ($x$ *communicates* with $y$) if and only if $x \rightsquigarrow y$ and $y \rightsquigarrow x$.

In this case, the relation is symmetric and transitive. As such, the relation $\leftrightsquigarrow$ forms an equivalence relation on $X$, and allows one to partition $X$ into disjoint subsets that are called *communicating classes* (or *equivalence classes*) (Chung, 1960).

**Definition 4.1.3.** Let $C(x)$ be the class containing $x \in X$. A state space $X$ consisting of a single class $C(x) = X$ (and the Markov chain operating on $X$) is said to be *irreducible*.

There are intimate relationships between the class structure of a coevolutionary Markov chain and the connectivity structure of the coevolutionary digraph that the chain operates on. One obvious connection is the notion of irreducibility.

---

[10]For example, the permutation matrix for a permutation vector $(2, 3, 1)$ is

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

**Lemma 4.1.4.** *A coevolutionary Markov chain is irreducible if it operates on an irreducible coevolutionary digraph.*

*Proof.* An irreducible digraph $D$ is strongly connected whereby for every $x, y \in V(D)$, there is an $(x, y)$-path and $(y, x)$-path. This implies that $x \leftrightarrow y$ for all $x, y \in X$ for a coevolutionary Markov chain operating on $X = V(D)$. There is only a single communicating class $C(x) = X$, in which case the coevolutionary Markov chain is irreducible. $\square$

**Definition 4.1.5.** States can be classified as *essential* or *inessential*. A state $x$ is essential if and only if for all $y$, $x \rightsquigarrow y$ implies $y \rightsquigarrow x$. Otherwise, $x$ is inessential when there is a $y$ such that $x \rightsquigarrow y$ but $y \not\rightsquigarrow x$.

The property that a state is essential or inessential is a *class property* (all the states in a communicating class share the same property). An essential state cannot lead to an inessential state. Crucially, an essential class is by definition *minimal closed* – it is closed in the sense that $\forall y \in C(x)$ $\mathbb{P}(y, C(x)) = 1$ and minimal in that it does not contain a proper subset that is closed (Chung, 1960). In other words, a coevolutionary Markov chain will be *absorbed* into an essential class upon entering it (i.e., it will be unable to leave it once it first enters it). This leads to the following lemma.

**Lemma 4.1.6.** *For a reducible coevolutionary Markov chain, there is only one absorbing class. The absorbing class consists of states corresponding to dominant solutions in the coevolutionary digraph.*

*Proof.* A coevolutionary Markov chain operating on a reducible coevolutionary digraph is reducible. By Lemma 2.7.8, a reducible coevolutionary semicomplete digraph will have a terminal strong component that dominates all other remaining vertices. States corresponding to the terminal strong component form an essential communication class $C(x) = \{y : x \leftrightarrow y\}$ that is also absorbing. $\square$

One consequence is that the state space can be partitioned into two disjoint subsets: (i) the absorbing class ($C$) of essential states and (ii) the class of inessential states ($B$). One can reindex the Markov transition matrix from essential to inessential states and obtain the following decomposition of $\boldsymbol{P}$ (Meyn and Tweedie, 2009)

$$
\boldsymbol{P} = \left( \begin{array}{c|ccc} & 0 & \cdots & 0 \\ C & \vdots & \ddots & \vdots \\ & 0 & \cdots & 0 \\ \hline & & B & \end{array} \right),
$$

which would allows us to identify and analyze the behaviors of the reducible coevolutionary Markov chain separately on the states in $C$ and $B$. We have the following for a reducible coevolutionary Markov chain.

**Lemma 4.1.7.** *Consider a reducible coevolutionary Markov chain with Markov transition matrix $\boldsymbol{P}$ having an absorbing class $C$. Let $\boldsymbol{P}_C$ be the matrix restricted to states in $C$. Then, there exists an irreducible coevolutionary Markov chain $\boldsymbol{\Phi}_C$ with state space restricted to $C$ and a Markov transition matrix given by $\boldsymbol{P}_C$.*

*Proof.* This is a direct consequence of Proposition 4.1.2 (page 79 in (Meyn and Tweedie, 2009)). $\square$

It remains then to investigate the behavior of the coevolutionary Markov chain on $B$ and the manner in which it gets absorbed to $C$. The next section will provide the means for which this can be done quantitatively.

## 4.2 Hitting Times

The analysis of the behaviors of the chain $\mathbf{\Phi}$ is often concerned with its distributions at certain *random times* as it evolves on $X$. Additional tools are needed for which these behaviors can be described in a precise and quantitative manner. One of the quantities refers to the random times with which $\mathbf{\Phi}$ visits certain states in $X$.

**Definition 4.2.1.** Let $\mathcal{A} \subset \mathcal{B}(X)$, where $\mathcal{B}(X)$ is the set of all subsets of $X$. The *first hitting time* of $\mathcal{A}$ is the random variable $\tau_{\mathcal{A}} : X \to \mathbb{N}_0 \cup \{\infty\}$, which is given by

$$\tau_{\mathcal{A}} = \inf\{t \geq 0 : \Phi_t \in \mathcal{A}\}. \tag{9}$$

The first hitting time of a state $x \in X$ is simply $\tau_x = \inf\{t \geq 0 : \Phi_t = x\}$. The first hitting time of an empty set $\varnothing$ is defined as $\infty$ by convention (Norris, 1998).

One can also define the first visit to $\mathcal{A}$ to only consider positive times $\tau_{\mathcal{A}}^+ = \inf\{t \geq 1 : \Phi_t \in \mathcal{A}\}$ (Meyn and Tweedie, 2009).

**Definition 4.2.2.** The *first return time* is defined as

$$\varrho_{\mathcal{A}} = \inf\{t \geq 1 : \Phi_t \in \mathcal{A}\}. \tag{10}$$

Obviously, $\varrho_{\mathcal{A}}$ is the case of $\tau_{\mathcal{A}}^+$ with $\Phi_0 \in \mathcal{A}$.[11] Another quantity relates to the number of visits to some states.

**Definition 4.2.3.** The *occupation time* $\eta_{\mathcal{A}}$ is defined as

$$\eta_{\mathcal{A}} = \sum_{t=1}^{\infty} \mathbb{I}\{\Phi_t \in \mathcal{A}\}, \tag{11}$$

where $\mathbb{I}$ is the indicator function whereby

$$\mathbb{I}\{\Phi_t \in \mathcal{A}\} = \begin{cases} 1 & \text{if } \Phi_t \in \mathcal{A} \\ 0 & \text{if } \Phi_t \notin \mathcal{A}. \end{cases}$$

The occupation time $\eta_{\mathcal{A}}$ counts the number of visits of a coevolutionary Markov chain $\mathbf{\Phi}$ to $\mathcal{A}$ after time zero. It is crucial in the following characterization of the set $\mathcal{A}$.

**Definition 4.2.4.** $\mathcal{A}$ is uniformly *transient* if there is a finite $\mathbb{E}_x(\eta_{\mathcal{A}})$ for all $x \in \mathcal{A}$. Otherwise, $\mathcal{A}$ is *recurrent* if $\mathbb{E}_x(\eta_{\mathcal{A}}) = \infty$ for all $x \in \mathcal{A}$.

This can be extended to individual states as well, i.e., a state is recurrent if $\mathbb{E}_x(\eta_x) = \infty$ (Meyn and Tweedie, 2009). This leads to the following lemma.

**Lemma 4.2.5.** *Consider an irreducible coevolutionary Markov chain $\mathbf{\Phi}$ on $X$. Every state $x \in X$ is recurrent.*

*Proof.* From Lemma 4.1.4, an irreducible $\mathbf{\Phi}$ operates on $X$ that coincides with an irreducible coevolutionary digraph $D = (V, A)$. Such a digraph is hamiltonian from Lemma 2.7.6, in which case there is a closed finite-length $(x, x)$-path for all $x \in V(D)$. This implies that $\mathbb{P}_x(\varrho_{\mathcal{A}} < \infty) = 1$ for all $x \in X$. Applying Proposition 8.1.3 in (Meyn and Tweedie, 2009), we have that $\mathbb{E}_x(\eta_x) = \infty$. As such, all $x$ is recurrent. $\qquad\square$

---

[11] First hitting and return times are *stopping times* in the sense that the event $\{\tau = t\}$ depends only on $\Phi_0, \Phi_1, \Phi_2, \dots, \Phi_t$ for $t = 0, 1, 2, \dots$, which follows from the *strong Markov property* (e.g., Theorem 1.4.2 in (Norris, 1998)). The Markov property is then said to hold for those random times. Furthermore, given that the coevolutionary Markov chain is discrete time, the strong Markov property always hold (by Proposition 3.4.6 in (Meyn and Tweedie, 2009)).

How about the case of a reducible coevolutionary Markov chain? Obviously, its state space can be partitioned into two classes – a closed absorbing class of recurrent states and another consisting of transient states only. As in the context of analysis of *Evolutionary Algorithms* (EAs) (He and Yao, 2003), studying the expected (mean) hitting time of an absorbing coevolutionary Markov chain $\boldsymbol{\Phi}$ is crucial. Unlike EAs, complete connectivities in the coevolutionary digraph means $\tau_{\mathcal{A}}^+ = 1$ although it is a rare event for $\boldsymbol{\Phi}$ to make a one-step walk to $\mathcal{A}$. The following gives a precise definition for this quantity (Norris, 1998).

**Definition 4.2.6.** The expected hitting time of $\mathcal{A}$ starting from $x \in X$ is

$$\mathbb{E}_x(\tau_{\mathcal{A}}) = \sum_{t < \infty} t \mathbb{P}_x(\tau_{\mathcal{A}} = t) + \infty \mathbb{P}_x(\tau_{\mathcal{A}} = \infty),$$

which is simplified to

$$\mathbb{E}_x(\tau_{\mathcal{A}}) = \sum_{t < \infty} t \mathbb{P}_x(\tau_{\mathcal{A}} = t) \tag{12}$$

for non-empty, reachable $\mathcal{A}$ (i.e. $\mathbb{P}_x(\tau_{\mathcal{A}} = \infty) = 0$). $\mathbb{P}_x(\tau_{\mathcal{A}} = t)$ is the probability that starting for state $x \in X$ that $\boldsymbol{\Phi}$ hits $\mathcal{A}$ at time $t$.

To compute $\mathbb{E}_x(\tau_{\mathcal{A}})$, one reindexes the states in the reducible Markov chain and make the appropriate simplification to produce the *canonical form* of its Markov transition matrix, $\boldsymbol{P}^*$ (Meyer, 2000):

$$\boldsymbol{P}^* = \left[ \begin{array}{cc} \mathbf{I} & \mathbf{0} \\ \mathbf{R} & \mathbf{Q} \end{array} \right].$$

$\mathbf{R}$ represents transition probabilities from transient states to the absorbing class. $\mathbf{Q}$ represents transition probabilities among transient states. $\mathbf{0}$ is a row vector of zeroes. The following summarizes our results.

**Lemma 4.2.7.** *Consider a reducible coevolutionary Markov chain with Markov transition matrix in its canonical form $\boldsymbol{P}^*$. Let $i \in \mathcal{Q}$ be the transient states. Let the column vector $\mathbf{h} = [h_i]_{i \in \mathcal{Q}}$ where $h_i = \mathbb{E}_i(\tau_{\mathcal{A}})$ is the expected hitting time starting from a transient state i to the absorbing class $\mathcal{A}$. Then,*

$$\mathbf{h} = (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}, \tag{13}$$

*where $\mathbf{1}$ is a column vector of ones.*

*Proof.* This is a direct consequence of Theorem 3.2 in (Iosifescu, 1980) for absorbing Markov chains. In our case, it follows from Lemma 4.1.6 that a reducible coevolutionary Markov chain is an absorbing Markov chain with one absorbing class. $\square$

For an absorbing Markov chain, the matrix $(\mathbf{I} - \mathbf{Q})^{-1}$ exists and is called the *fundamental matrix* of the Markov transition matrix $\boldsymbol{P}$. One can perform direct computation of the expected hitting time of any reducible coevolutionary Markov chain by first computing its fundamental matrix (Theorem 11.4 in (Grinstead and Snell, 1997)):

$$(\mathbf{I} - \mathbf{Q})^{-1} = \mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \mathbf{Q}^3 + \cdots.$$

We note that this series converges given that $\mathbf{Q}^n \to 0$ as $n \to \infty$ (Theorem 11.3 in (Grinstead and Snell, 1997)).

## 4.3 Invariant Measures - Stationary Distribution

This section is concerned with characterizing the coevolutionary Markov chain $\boldsymbol{\Phi}$ in the context of the long-time properties of $\boldsymbol{\Phi}$. This is relevant for irreducible coevolutionary Markov chains as well as reducible chains (after the chain enters the absorbing class). We first introduce the notion of an *invariant measure* (Norris, 1998).

**Definition 4.3.1.** Let the row vector $\boldsymbol{\pi} = (\pi_x : x \in X)$ be a probability measure on $X$ and $\boldsymbol{P}$ be the Markov transition matrix. Then, $\boldsymbol{\pi}$ is said to be *invariant* if

$$\boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi}.$$

A $\boldsymbol{\pi}$ that is defined on $X$ and satisfies the invariant property would represent the long-term limiting distribution of the coevolutionary Markov chain. In addition, such a distribution can be described in the following equivalent terms:

(i) *Stationary* - If $(\Phi_n)_{n \geq 0}$ is a coevolutionary Markov chain with $\boldsymbol{\mu} = \boldsymbol{\pi}$ and $\boldsymbol{P}$, then $(\Phi_{m+n})_{n \geq 0}$ having $\mathbb{P}(\Phi_m = x) = (\boldsymbol{\pi}\boldsymbol{P}^m)_x = \pi_x$ for all $x \in X$ is also a coevolutionary Markov chain with $\boldsymbol{\mu} = \boldsymbol{\pi}$ and $\boldsymbol{P}$ (Theorem 1.7.1 in (Norris, 1998)).

(ii) *Equilibrium* - For all $x \in X$,[12] $\mathbb{P}^n(x, y) \to \pi_y$ as $n \to \infty$ for all $y \in X$ (Theorem 1.7.2 in (Norris, 1998)).

It is known that for an irreducible Markov chain $\boldsymbol{\Phi}$ on countable state space $X$, its stationary distribution can be calculated from the *expected return times* $\mathbb{E}_x(\varrho_x)$ (Levin et al., 2009). The general intuition is that if one breaks up the trajectory of $\boldsymbol{\Phi}$ into identically distributed segments through visits to and from some arbitrary states, the average fraction of time that is spent on a state $x \in X$ by each segment would be similar to the long-term fraction of time spent on $x$ by $\boldsymbol{\Phi}$. Furthermore, starting from any arbitrary distribution $\boldsymbol{\mu}$, the chain would converge to its stationary distribution in the long run (Norris, 1998; Levin et al., 2009). For coevolutionary Markov chains, we have the following results.

**Lemma 4.3.2.** *Consider an irreducible coevolutionary Markov chain with Markov transition matrix $\boldsymbol{P}$. Then, there is a unique invariant probability distribution $\boldsymbol{\pi} = (\pi_x : x \in X)$ on $X$ where $0 \leq \pi_x \leq 1$ and $\sum_{x \in X} \pi_x = 1$ such that $\boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi}$. Furthermore, $\pi_x = 1/\mathbb{E}_x(\varrho_x)$.*

*Proof.* This is a direct consequence of Theorems 1.7.7 in (Norris, 1998) whereby an irreducible and positive recurrent Markov chain on a finite state space has a unique invariant probability measure $\boldsymbol{\pi} = (\pi_x : x \in X)$ on $X$ with $\pi_x = 1/\mathbb{E}_x(\varrho_x)$. A state $x$ is positive recurrent if its expected return time $\mathbb{E}_x(\varrho_x)$ is finite. By Lemma 4.1.4, obviously all $x \in X$s are positive recurrent since $\mathbb{E}_x(\varrho_x)$ is finite. $\square$

**Lemma 4.3.3.** *Consider a reducible coevolutionary Markov chain with Markov transition matrix $\boldsymbol{P}$. For all inessential states $y \in X$, $\pi_y = 0$. Then, there is a unique invariant probability distribution $\boldsymbol{\pi}$ concentrated on the absorbing class $C$ of essential states in $X$.*

*Proof.* Let $B$ be the set of all inessential states $y \in X$. The first statement on $\pi_y = 0$ for all inessential states $y$ is a direct consequence of Proposition 1.25 (page 16 in (Levin et al., 2009)). This implies that the probability distribution is concentrated on the remaining essential states of the absorbing class $C$. From Lemma 4.1.7, we can study the irreducible

---

[12] Given the complete connectivities of coevolutionary digraphs.

coevolutionary Markov chain with $\boldsymbol{P}_C$ separately and obtain $\boldsymbol{\pi}_C = (\pi_x : x \in C(x))$ as a result of Lemma 4.3.2. Since $X = C \cup B$ and $C \cap B = \varnothing$, then $\boldsymbol{\pi} = (\boldsymbol{\pi}_C, \boldsymbol{\pi}_B)$ whereby all elements in $\boldsymbol{\pi}_B$ are zero. □

**Lemma 4.3.4.** *Consider an irreducible coevolutionary Markov chain $\boldsymbol{\Phi}$ with Markov transition matrix $\boldsymbol{P}$ and any arbitrary initial distribution $\boldsymbol{\mu}$. Let $x, y \in X$. Let $\boldsymbol{\pi}$ be the unique invariant probability distribution of $\boldsymbol{\Phi}$. Since $\boldsymbol{P}$ is irreducible and aperiodic,*

$$\mathbb{P}(\Phi_n = y) \to \pi_y \text{ as } n \to \infty \text{ for all } y.$$

*Furthermore,*

$$\mathbb{P}^n(x, y) \to \pi_y \text{ as } n \to \infty \text{ for all } x, y.$$

*Proof.* This is a direct consequence of the *Convergence to Equilibrium Theorem* (given as Theorem 1.8.3 in (Norris, 1998)). We just need to show that an irreducible coevolutionary Markov chain $\boldsymbol{\Phi}$ is aperiodic. By Lemma 2.7.6, irreducible coevolutionary digraphs are vertex pancyclic, in which case the greatest common divisor to the cycle lengths is one. As such, $\boldsymbol{\Phi}$ is aperiodic. □

## 4.4 Case Studies on Quantitative Measures of Coevolutionary Markov Chains

We first provide analytic expression for the expected hitting time of specific coevolutionary Markov chains. This case study considers both the random walk and (1+1)CEA on transitive coevolutionary tournaments. We note that the Markov transition matrices given by Equations 5 and 6 are in the form of an upper triangle matrix. It requires relabelling the indices in $\boldsymbol{P}$ so that it becomes a lower triangle matrix. For these two cases, $\mathbf{Q}$s are also in the form of a lower triangle matrix. Crucially, each of the $n \times n$ matrices $(\mathbf{I} - \mathbf{Q})$ has a $rank(\mathbf{I} - \mathbf{Q}) = n$ so its inverse exists (Meyer, 2000).

It is straightforward to obtain the following for the random walk on labelled transitive coevolutionary tournament of order $n + 1$

$$
\mathbf{I} - \mathbf{Q} = \begin{pmatrix}
1 & 0 & 0 & \dots & 0 & 0 \\
-\frac{1}{2} & 1 & 0 & \dots & 0 & 0 \\
-\frac{1}{3} & -\frac{1}{3} & \ddots & \ddots & \vdots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
-\frac{1}{n-1} & -\frac{1}{n-1} & \ddots & -\frac{1}{n-1} & 1 & 0 \\
-\frac{1}{n} & -\frac{1}{n} & \dots & -\frac{1}{n} & -\frac{1}{n} & 1
\end{pmatrix}. \tag{14}
$$

It is possible to exploit the inductive structure of $\boldsymbol{P}$. We apply the *Gauss-Jordan* method to obtain the $m \times m$ fundamental matrix, and then use this result to partially fill the $m \times m$ submatrix of the $(m+1) \times (m+1)$ fundamental matrix. Starting from $m = 2$ and applying this iteratively, we obtain the following form:

$$
(\mathbf{I} - \mathbf{Q})^{-1} = \begin{pmatrix}
1 & 0 & 0 & \dots & 0 & 0 \\
\frac{1}{2} & 1 & 0 & \dots & 0 & 0 \\
\frac{1}{2} & \frac{1}{3} & \ddots & \ddots & \vdots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
\frac{1}{2} & \frac{1}{3} & \ddots & \frac{1}{n-1} & 1 & 0 \\
\frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{n-1} & \frac{1}{n} & 1
\end{pmatrix}. \tag{15}
$$

To demonstrate the inductive structure, note that as the fundamental matrix increases in size, a new row is added to the bottom and a new column is added to the left. Observe that the element in the first column and last row of $(\mathbf{I} - \mathbf{Q})^{-1}$ always add to $1/2$. For $2 \times 2$, it is obviously $1/2$ due to Gauss-Jordan operations. Increasing the matrix in size,

$$3 \times 3: \qquad \frac{1}{3} \times \frac{1}{2} + \frac{1}{3} \times 1 \;=\; \frac{1}{3} \times \left(\frac{3}{2}\right)$$
$$= \frac{1}{2}$$
$$4 \times 4: \qquad \frac{1}{4} \times \frac{1}{2} + \frac{1}{4} \times \frac{1}{2} + \frac{1}{4} \times 1 \;=\; \frac{1}{4} \times \left(\frac{4}{2}\right)$$
$$= \frac{1}{2}$$
$$5 \times 5: \quad \frac{1}{5} \times \frac{1}{2} + \frac{1}{5} \times \frac{1}{2} + \frac{1}{5} \times \frac{1}{2} + \frac{1}{5} \times 1 \;=\; \frac{1}{5} \times \left(\frac{5}{2}\right)$$
$$= \frac{1}{2}.$$

For $(n+1) \times (n+1)$, we note that

$$\overbrace{\frac{1}{n+1} \times \frac{1}{2} + \frac{1}{n+1} \times \frac{1}{2} + \cdots + \frac{1}{n+1} \times \frac{1}{2}}^{n-1} + \frac{1}{n+1} \times 1$$
$$= \frac{1}{n+1} \times \left(\frac{n+1}{2}\right)$$
$$= \frac{1}{2}.$$

It is then a simple but laborious matter to compute and verify that $(\mathbf{I}-\mathbf{Q})(\mathbf{I}-\mathbf{Q})^{-1} = \mathbf{I}$. We apply Lemma 4.2.7 to calculate the expected hitting time and then obtain

$$\mathbf{h} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = \left(1, \frac{1}{2} + 1, \frac{1}{2} + \frac{1}{3} + 1, \ldots, \sum_{i=1}^{n} \frac{1}{n}\right)^{T}. \qquad (16)$$

For the case of the (1+1)CEA, direct calculation reveals:

$$\mathbf{I} - \mathbf{Q} = \begin{pmatrix} \frac{1}{n} & 0 & 0 & \ldots & 0 & 0 \\ -\frac{1}{n} & \frac{2}{n} & 0 & \ldots & 0 & 0 \\ -\frac{1}{n} & -\frac{1}{n} & \frac{3}{n} & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \ddots & -\frac{1}{n} & \frac{n-1}{n} & 0 \\ -\frac{1}{n} & -\frac{1}{n} & \ldots & -\frac{1}{n} & -\frac{1}{n} & 1 \end{pmatrix}, \qquad (17)$$

$$(\mathbf{I} - \mathbf{Q})^{-1} = \begin{pmatrix} n & 0 & 0 & \ldots & 0 & 0 \\ \frac{n}{2} & \frac{n}{2} & 0 & \ldots & 0 & 0 \\ \frac{n}{2} & \frac{n}{2(3)} & \frac{n}{3} & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{n}{2} & \frac{n}{2(3)} & \ddots & \frac{n}{(n-2)(n-1)} & \frac{n}{n-1} & 0 \\ \frac{n}{2} & \frac{n}{2(3)} & \ldots & \frac{n}{(n-2)(n-1)} & \frac{n}{(n-1)(n)} & 1 \end{pmatrix}, \qquad (18)$$

$$\mathbf{h} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1} = (n, n, n, \ldots, n)^{T}. \qquad (19)$$

In this second part of our case study, we calculate the stationary distribution of a specific class of coevolutionary Markov chain – those operating on regular tournaments. We can make use of structural results from digraph theory to obtain a general result on the stationary distribution of regular coevolutionary Markov chains.

**Lemma 4.4.1.** *Consider a coevolutionary Markov chain operating on a regular coevolutionary tournament $T(V_{S(n)}, A) \in \mathcal{T}(V_{S(n)})$ of odd order $n \geq (3 + 2m)$, $m = 0, 1, 2, \ldots$. This regular coevolutionary Markov chain has a unique invariant probability measure $\boldsymbol{\pi} = (\pi_x : x \in X)$ on $X = V_{S(n)}$ whereby $\pi_x = 1/n$ for all $x$.*

*Proof.* Every regular coevolutionary tournament is *arc-pancyclic*, which is a direct consequence of Corollary 6.6.4 in (Bang-Jensen and Gutin, 2009). An arc-pancyclic coevolutionary tournament is arc-$k$-cyclic for all $k = \{3, 4, 5, \ldots, n\}$, i.e., each arc of $T(V_{S(n)}, A)$ is contained in a cycle for each length of $k = \{3, 4, 5, \ldots, n\}$. Since $d_D^+(x) = d_D^-(x) = (n-1)/2$ for all $x$ in a regular $T(V_{S(n)}, A)$, this means that there is an equal number of arcs of various lengths $k = \{3, 4, 5, \ldots, n\}$ that start and end at each $x \in X$. Any random walk on $T(V_{S(n)}, A)$ would thus spend equal fraction of time for each $x$, in which case $\pi_x = 1/n$ for a tournament with $n$ vertices. □

One immediate consequence of Lemma 4.4.1 is that any coevolutionary process (random walk or (1+1)CEA for example) can be defined on a regular tournament of order $n$ and that $\boldsymbol{\pi} = (1/n, 1/n, 1/n, \ldots, 1/n)$ since the cycles that start from and return to each vertex are the same. We can also apply this result with Lemma 4.3.3 to calculate the stationary distribution of a reducible coevolutionary Markov chain having an absorbing class that is essentially a regular subtournament.

### 4.4.1 Absorbing Coevolutionary Markov Chains

Earlier result for (1+1)CEA with uniform expected hitting time for any starting vertex indicate unique properties (the diagonal elements in the probability transition matrix), which we can generalize further with the following result.

**Theorem 4.4.2.** *Consider a (1+1)CEA operating on a reducible coevolutionary tournament $T(V_{S(n+n_d)}, A) \in \mathcal{T}(V_{S(n+n_d)})$. Let $V_{S(n+n_d)}$ consists of two disjoint subsets $V_{S(n)}^1$ and $V_{S(n_d)}^2$, whereby $V_{S(n)}^1 \cap V_{S(n_d)}^2 = \varnothing$. Let $V_{S(n_d)}^2 \Leftarrow V_{S(n)}^1$. Let the coevolutionary Markov chain be defined with Markov transition matrix in its canonical form $\boldsymbol{P}^*$. Let $i \in \mathcal{Q}$ be the transient states and that $|\mathcal{Q}| = n$. Let the column vector $\mathbf{h} = [h_i]_{i \in \mathcal{Q}}$ where $h_i = \mathbb{E}_i(\tau_{\mathcal{A}})$ is the expected hitting time starting from a transient state $i$ to the absorbing class $\mathcal{A} = V_{S(n_d)}^2$. Then, $\mathbf{h} = [h_i]_{i \in \mathcal{Q}} = [n, n, n, \ldots, n]^T$.*

*Proof.* It follows from Lemma 2.7.8 that the absorbing class $\mathcal{A} = V_{S(n_d)}^2$ must be a strong component. One can perform a contraction operation on $V_{S(n_d)}^2$ and replace it with a single vertex and remove all parallel arcs. The resulting coevolutionary tournament is $T(V_{S(n+1)}, A)$. The absorbing class $\mathcal{A}$ is now a single absorbing state and that

$$\boldsymbol{P}^* = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{R} & \mathbf{Q} \end{bmatrix},$$

where $\mathbf{R}$ is a single column of $n$ elements, $\mathbf{Q}$ is a $n \times n$ matrix, and $\mathbf{0}$ is a single row of $n$ zeroes.

For every transient state $i \in \mathcal{Q}$, note that the corresponding vertex $i$ in the coevolutionary tournament $T(V_{S(n+1)}, A)$ has the same degree $d_D(i) = d_D^+(i) + d_D^-(i) = (n+1) - 1 = n$. Let $\mathbf{R} = (r_i : i = 1, 2, 3, \ldots, n)$ and $\mathbf{Q} = (q_{ij} : i, j = 1, 2, 3, \ldots, n)$. Given our definition of (1+1)CEA (as in Algorithm 2), for each $i$: (1) $q_{ii} = \frac{d_D^-(i)}{d_D(i)}$ and (2) $(\sum_{j \in N_D^+(i) \smallsetminus x_a} q_{ij}) + r_i = \frac{d_D^+(i)}{d_D(i)}$. The one-step transition probability from $i$ to the single absorbing state $x_a$ is given by $r_i$. Since each one-step transition from $i$ to its

outneighbour $j$ occurs with equal probability, then $\forall j \in N_D^+(i)$ $q_{ij} = \frac{1}{d_D(i)}$ and $r_i = \frac{1}{d_D(i)}$. Note that $\forall j \in N_D^-(i)$ $q_{ij} = 0$. As such, each row of $\mathbf{R1} + \mathbf{Q1}$ sums to one, i.e., $\forall i$ $(r_i + \sum_{j=1}^n q_{ij}) = \frac{d_D(i)}{d_D(i)} = 1$.

Next, we need to show that the corresponding row sums of $(\mathbf{I} - \mathbf{Q})^{-1} = \mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \mathbf{Q}^3 + \cdots$ follow a geometric series. We need to prove that $\mathbf{Q}_{n \times n}^m \mathbf{1} = [(\frac{n-1}{n})^m, (\frac{n-1}{n})^m, (\frac{n-1}{n})^m, \ldots, (\frac{n-1}{n})^m]^T$. This will give the expected hitting times calculated from the geometric series $1 + \frac{n-1}{n} + (\frac{n-1}{n})^2 + (\frac{n-1}{n})^3 + \cdots = 1/(1 - \frac{n-1}{n}) = n$. We do this by induction. Let the column vector $\mathbf{Q}_{n \times n}^m \mathbf{1} = [Q^m]_{i=1}^n$, with $[Q^m]_i$ representing the $i$th element of $[Q^m]_{i=1}^n$. The induction hypothesis is that $\mathbf{Q}_{n \times n}^{m+1} \mathbf{1} = [Q^m Q^1]_{i=1}^n$.

Basis $(m = 0)$: By definition $\mathbf{Q}_{n \times n}^0 = \mathbf{I}_{n \times n}$. Then, $\mathbf{Q}_{n \times n}^0 \mathbf{1} = \mathbf{I}_{n \times n} \mathbf{1} = \mathbf{1}$.

Inductive step: We need to show that $\mathbf{Q}_{n \times n}^{m+1} \mathbf{1} = [Q^m Q^1]_{i=1}^n$. We first note that $\mathbf{Q}^{m+1} \mathbf{1} = \mathbf{Q}^m \mathbf{Q}^1 \mathbf{1}$. Let $\mathbf{Q}_{n \times n}^m = (x_{ij} : i, j = 1, 2, 3, \ldots, n)$ and $\mathbf{Q}_{n \times n}^1 = (y_{ij} : i, j = 1, 2, 3, \ldots, n)$.. Let $\mathbf{Q}_{n \times n}^m \mathbf{Q}_{n \times n}^1 = (z_{ij} : i, j = 1, 2, 3, \ldots, n)$. $Q_i^{m+1} = \sum_{j=1}^n z_{ij}$. Since $\forall i$ $r_i = \frac{1}{d_D(i)} = \frac{1}{n}$, then $\forall i$ $Q_i^1 = 1 - \frac{1}{n} = \frac{n-1}{n}$.

For $Q_1^{m+1} = \sum_{j=1}^n z_{1j}$, we have

$$
\begin{array}{llllllllll}
z_{11} & = & x_{11}y_{11} & + & x_{12}y_{21} & + & x_{13}y_{31} & + & \cdots & + & x_{1n}y_{n1} \\
z_{12} & = & x_{11}y_{12} & + & x_{12}y_{22} & + & x_{13}y_{32} & + & \cdots & + & x_{1n}y_{n2} \\
z_{13} & = & x_{11}y_{13} & + & x_{12}y_{23} & + & x_{13}y_{33} & + & \cdots & + & x_{1n}y_{n3} \\
\vdots & \vdots & & & & \vdots & & & & & \\
z_{1n} & = & x_{11}y_{1n} & + & x_{12}y_{2n} & + & x_{13}y_{3n} & + & \cdots & + & x_{1n}y_{nn}.
\end{array}
$$

We can factor the terms out such that

$$
\begin{aligned}
Q_1^{m+1} \ =\ & x_{11}(y_{11} + y_{12} + y_{13} + \cdots + y_{1n}) \\
& + x_{12}(y_{21} + y_{22} + y_{23} + \cdots + y_{2n}) \\
& + x_{13}(y_{31} + y_{32} + y_{33} + \cdots + y_{3n}) \\
& + \\
& \vdots \\
& + x_{1n}(y_{n1} + y_{n2} + y_{n3} + \cdots + y_{nn}) \\
=\ & x_{11}Q_1^1 + x_{12}Q_2^1 + x_{13}Q_3^1 + \cdots + x_{1n}Q_n^1 \\
=\ & (x_{11} + x_{12} + x_{13} + \cdots + x_{1n})\frac{n-1}{n} \\
=\ & Q_1^m \frac{n-1}{n} \\
=\ & Q_1^m Q_1^1
\end{aligned}
$$

since each $Q_i^1 = \frac{n-1}{n}$. Similar calculations can be made for all other rows. As such, $\forall i$ $Q_i^{m+1} = Q_i^m Q_i^1$, in which case, $\mathbf{Q}_{n \times n}^{m+1} \mathbf{1} = [Q^m Q^1]_{i=1}^n$.

From our induction hypothesis, it can be seen that $\mathbf{Q}_{n \times n}^m \mathbf{1} = [Q^m]_{i=1}^n = [(\frac{n}{n+1})^m]_{i=1}^n$. Each row sum of $\mathbf{h} = (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1} = (\mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \mathbf{Q}^3 + \cdots) \mathbf{1}$ then follows the geometric series $1 + \frac{n-1}{n} + (\frac{n-1}{n})^2 + (\frac{n-1}{n})^3 + \cdots = 1/(1 - \frac{n-1}{n}) = n$. Thus, $\mathbf{h} = [h_i]_{i \in \mathcal{Q}} = [n, n, n, \ldots, n]^T$. $\square$

Example: Note that $V_{S(n)}^1$ itself can be a single strong component of $T(V_{S(n+n_d)}, A)$ of any odd order. For the case where it is a strong component of order one, direct calculations can be made. Consider the $(1+1)$CEA on an unlabelled coevolutionary tournament $T(V_{S(n+1)}, A)$ with a single dominant vertex and a single dominated strong component that consists of a regular tournament of odd order $n \geq (3+2k)$, $k = 0, 1, 2, \ldots$. We then have the following:

$$\boldsymbol{P^*} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{n} & \frac{1}{2}\frac{n-1}{n} & 0 & \frac{1}{n} & 0 & \cdots & 0 & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \frac{1}{2}\frac{n-1}{n} & 0 & \frac{1}{n} & \cdots & \frac{1}{n} & 0 \\ \frac{1}{n} & 0 & \frac{1}{n} & \frac{1}{2}\frac{n-1}{n} & 0 & \cdots & 0 & \frac{1}{n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & 0 & \frac{1}{n} & 0 & \cdots & \frac{1}{2}\frac{n-1}{n} & 0 \\ \frac{1}{n} & 0 & \frac{1}{n} & 0 & \frac{1}{n} & \cdots & \frac{1}{n} & \frac{1}{2}\frac{n-1}{n} \end{pmatrix}.$$

In particular, we can write $\mathbf{Q}$ in the matrix form with $a = \frac{1}{2}\frac{n-1}{n}$ and $b = \frac{1}{n}$:

$$\mathbf{Q} = \begin{pmatrix} a & 0 & b & 0 & \cdots & 0 & b \\ b & a & 0 & b & \cdots & b & 0 \\ 0 & b & a & 0 & \cdots & 0 & b \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ b & 0 & b & 0 & \cdots & a & 0 \\ 0 & b & 0 & b & \cdots & b & a \end{pmatrix}.$$

Note that the $(i+1)$th row is a right-shift of the $i$th row of $\mathbf{Q}$. Furthermore, each of the $i$th row sum $Q_i$ has the same value and is given by $a + \left(\frac{n-1}{2}\right)b = \frac{1}{2}\frac{n-1}{n} + \left(\frac{n-1}{2}\right)\frac{1}{n} = \frac{n-1}{n}$. This follows from the definition of a regular tournament whereby each vertex $i$ has $d_D^+(i) = d_D^-(i)$. $\mathbf{Q}_{n\times n}\mathbf{1} = \left[\frac{n-1}{n}, \ldots, \frac{n-1}{n}\right]^T$. Obviously, one can obtain the same result by taking $\mathbf{1} - \mathbf{R1}$ since $\mathbf{R} = \left[\frac{1}{n}, \ldots, \frac{1}{n}\right]^T$ and that $(\mathbf{R1} + \mathbf{Q1}) = \mathbf{1}$.

The procedure to show that the row sums of $(\mathbf{I} - \mathbf{Q})^{-1} = \mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \mathbf{Q}^3 + \cdots$ follow a geometric series is the same as in the Theorem 4.4.2. It can be shown that $\mathbf{Q}_{n\times n}^m\mathbf{1} = [Q^m]_{i=1}^n = \left[\left(\frac{n-1}{n}\right)^m\right]_{i=1}^n$. Each element of $\mathbf{h} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{1}$ then follows the geometric series $1 + \frac{n-1}{n} + \left(\frac{n-1}{n}\right)^2 + \left(\frac{n-1}{n}\right)^3 + \cdots = 1/\left(1 - \frac{n-1}{n}\right) = n$, in which case, $\mathbf{h} = [h_i]_{i\in\mathcal{Q}} = [n, n, n, \ldots, n]^T$.

To verify, note that

$$\mathbf{I} - \mathbf{Q} = \begin{pmatrix} 1-a & 0 & -b & 0 & \cdots & 0 & -b \\ -b & 1-a & 0 & -b & \cdots & -b & 0 \\ 0 & -b & 1-a & 0 & \cdots & 0 & -b \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ -b & 0 & -b & 0 & \cdots & 1-a & 0 \\ 0 & -b & 0 & -b & \cdots & -b & 1-a \end{pmatrix}.$$

Each of the $i$th element of $(\mathbf{I} - \mathbf{Q})\mathbf{h}$ is exactly one, i.e., $(1-a)n + \left(\frac{n-1}{2}\right)(-b)n = \left(1 - \frac{1}{2}\frac{n-1}{n}\right)n + \left(\frac{n-1}{2}\right)\left(-\frac{1}{n}\right)n = \frac{2n-n+1}{2} - \frac{n-1}{2} = \frac{2}{2} = 1$.

## 4.5 Generating Random Coevolutionary Tournaments

We conclude this section with a computational study on generating random coevolutionary tournaments in a controlled manner. Potential applications include using random tournaments as benchmarks to study methodologies estimating quantities of interest in

more complex and realistic coevolutionary Markov chains. We first note that the simple approach of randomly orienting edges of the underlying graphs would usually generate irreducible random tournaments (Moon, 1968). There is more than 50% chance to generate an irreducible random tournament of order five, which increases to more than 99.9% for random tournaments of order 16! Beyond that, generated random tournaments would almost certainly be irreducible. Another approach is to first generate all valid score sequences (Hemasinha, 2003), which are used to reconstruct the corresponding tournaments (Moon, 1968).

---

**Algorithm 3** Growth-based algorithm to generate random tournaments $T(V_{S(n)}, A)$

---

1: **procedure** GROW($T(V, A), n_{\text{init}}, n_{\text{final}}$)
2:   $n := 0$
3:   **Initialize $T(V, A)$ of order $n_{\text{init}}$**
4:   $n := n_{\text{init}}$
5:   **Grow Process**
6:   **while $n \leq n_{\text{final}}$ do**
7:     $V' := V,\ A' := A$
8:     $T(V', A')$ indexed according to its score sequence $(s_1, s_2, s_3, \ldots, s_n)$.
9:     Compute attachment probabilities for $V_{S(n)}$, $\{\mathbb{P}_i^-\}_1^n$.
10:     **Add new vertex $v_{n+1}$ to $T(V_{S(n)}, A)$**
11:     $V := V' \cup v_{n+1}$
12:     **for all $v_i \in V'$ do**
13:       $A := A' \cup (v_{n+1}, v_i)$ with probability $\mathbb{P}_i^-$
14:     **end for**                  $\triangleright A := A' \cup (v_i, v_{n+1})$ with probability $\mathbb{P}_i^+ = 1 - \mathbb{P}_i^-$
15:     $n := n + 1$
16:   **end while**
17: **end procedure**

---

We will now propose a methodology for generating a hierarchy of random coevolutionary tournaments where a control parameter allows us to move between two complexity extremes – irreducible tournaments and tournaments with prominent transitive structures. The method is motivated by network growth methodologies via preferential attachment in the area of complex networks (e.g. (Rudolf et al., 2012)). Algorithm 3 describes a general implementation. The tournament can be seeded with an initial, small-sized tournament in a random or arbitrary manner (e.g., acyclic ordering). A new node (vertex) is added to the existing tournament of order $n$ at each iteration $n+1$. This requires orienting $n + 1$ edges from the new node to the tournament. This orientation is done randomly and independently but can be biased in a controlled manner towards the existing vertices with higher scores. This requires computing attachment probabilities:

$$\mathbb{P}_i^- = \frac{1}{1 + e^{-\beta(x_i - \overline{x})}}, \tag{20}$$

where $\{x_i\}_1^n$ are the relevant statistics of the tournament. Parameter $\beta > 0$ controls the steepness of the sigmoid link function of the Bernoulli distribution. A lower value of $\beta$ reduces the steepness of the function with $\mathbb{P}_i^- \to 1/2$ for $\beta \to 0$ (which is the same as the case of random orientation discussed above).

This setting can be motivated by statistical physics. Each node of the existing tournament has an *'energy'* related to its score (in-degree). The new incoming node $n+1$ will be more likely to orient edges towards high energy node in the existing tournament.

Parameter $\beta$ plays the role of inverse temperature. For high temperature regimes, the energy of nodes begins to equalize and the probability of edge orientation is the same for any node in the existing tournament. In contrast, by cooling the system (increasing $\beta$), orientation of edges from the new node $n+1$ towards high energy nodes of the existing tournament becomes more persistent.

We have performed initial experiments to determine appropriate choices for the parameters involved. We initialize (seed) the graph growth with transitive tournaments. For the statistics $\{x_i\}_1^n$, we use the corresponding node ranks $(i)_1^n$ with $\overline{x} = n/2$, $\beta \in [0, 4]$. Tournaments of order $n_{\text{final}} = 1000$ are generated from initial ones of order $n_{\text{init}} \in \{10, 100\}$. Each experiment set is repeated 100 times, thus generating 100 tournaments randomly and independently. The results are summarized in Figure 3. Obviously, the number of strong components are higher when the initial tournament seeds are larger as shown in Figure 3(b).

We next introduce into our 1000-node tournaments an additional single dominant (1001-th) node that has all the edges from the original 1000 vertices oriented towards it. This construction makes the tournaments reducible and so we can compute the expected hitting time from the *weakest* node $x_1$ (with the lowest score) to the dominant (absorbing) node $x_{1001}$. We now study the impact of increasing transitive structures, measured by the number of strong components (reflected in the complexity (Landau's) Index $\nu$ (Equation 1)), on dynamics of the coevolutionary process, measured by the expected hitting times of the dominant node.

First, we show results for the experiment that uses the standard random walk (Algorithm 1) with initial seed tournament of order 10. We can clearly see signatures of higher and lower temperature regimes in Figure 4(a). In the high temperature regime ($\beta \approx [0, 0.001]$), all the base 1000-node tournaments are irreducible, hence there is only one strong component. Together with the dominant 1001-th node, there are two strong components. There is little variability of hitting times (around 492–497), which reflects limited variability of the internal structure of the irreducible 1000-node base tournaments. For comparison, the expected hitting time is exactly 500 for a regular 999-node base tournament. In the low temperature regime ($\beta \approx [1, 4]$), prominent transitive structure dominates the base tournaments and the hitting times range in $[7.486, 10.659]$. For comparison, the expected hitting times would be slightly less than 10 for a vertex pancyclic 1000-node base tournament (which has the longest transitive structure) and around 7.4855 for a transitive 1000-node base tournament.

As in most complex systems, in between these two temperature regimes sits an intermediate temperature regime that exhibits the most interesting behaviour. Full spectrum of possible values of Landau's index and expected hitting times can be observed. Indeed, this is the case where transitive structures co-exist with a wide range of hitting times, reflecting different degrees of cycle structures in the strong components. Figure 4 illustrates the relationship between the (inverse) temperature $\beta$ and the hitting times of the dominant 1001-th node. As expected, while extremely high temperatures lead to irreducible base tournaments with intricate cycle structures and long hitting times, very low temperature regimes lead to base tournaments with prominent trivial transitive structures, yielding minimal hitting times. It is also natural to expect that, on average, the expected hitting times will decrease with decreasing temperature, since lower temperatures induce more transitive structures in the base tournaments. The most 'complex' and interesting regimes lie in between these two extremes at $\beta \approx (0.005, 0.5)$.

Theorem 4.4.2 will indicate that the expected hitting times for (1+1)CEA operating on these generated tournaments to be the same as their orders (1001 in this case),
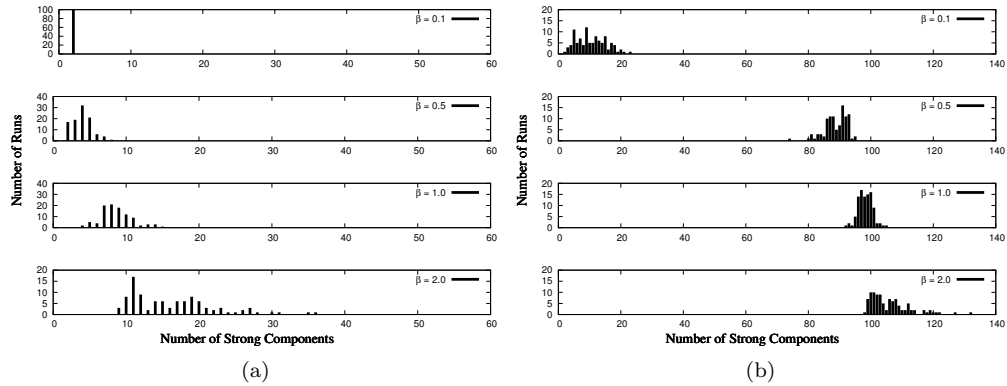
Figure 3: Histograms of number of runs generating random tournaments having specific number of strong components (Landau's Index). Corresponding experiment sets where initial transitive tournaments are of order (a) 10 and (b) 100.
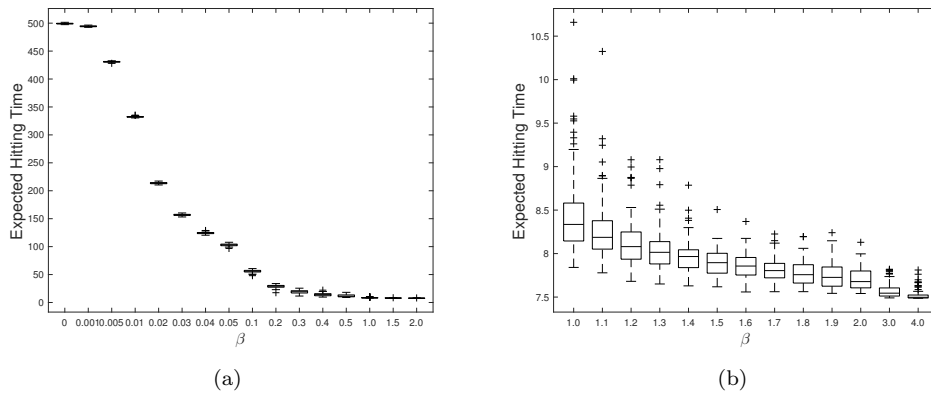


Figure 4: Boxplots with centerline representing the median, the two edges corresponding to the 25th and 75th percentiles, and outliers are plotted as '+' using $\beta$ set to a range in (a) $[0, 2]$ and (b) $[1, 4]$.

which are verified by repeating the experiment with (1+1)CEA. Computations for the expected hitting times are dependent on the probability transition matrix associated with the coevolutionary Markov chain under study. Obviously, different constructions of coevolutionary processes – standard random walk against (1+1)CEA – operating on the same tournaments can lead to different outcomes in the expected hitting time analysis. In contrast, a digraph-theoretic-based approach such as the Landau's index is unaffected by the specific constructions of coevolutionary processes. There are potential ramifications on the construction of sampling-based indices (e.g. hitting times) for analysis of structures in coevolutionary tournaments. Our controlled computational study suggests that more information related to the structures of the tournaments under study may need to be incorporated into those indices so that they are able to uncover those relevant structures. For example, a standard random walk incorporates information of connectivity structures (i.e. out-neighbourhoods of vertices) directly in contrast to the

(1+1)CEA.

## 5 Discussion and Conclusion

In this study, we have addressed the issue of studying rich cycle structures associated with coevolutionary systems that use pairwise-comparison-based fitness measures. We have developed a framework that fully captures the underlying structures in coevolutionary problems and represent them as cycle structures in digraphs so that we can begin to uncover their effects on coevolution. The digraph representation allows us to give precise characterizations of cycles of different lengths (from 3-cycles to Hamilton cycles that go through every vertex in the digraph) that can occur in coevolutionary problems.

Although these cycle structures appear complex at first given the variety of cycles of different lengths and the manner they can be oriented in digraphs, our study has revealed that coevolutionary digraphs admit a simple qualitative characterization – the corresponding coevolutionary problem is either solvable or not. This has a significant implication on coevolution when it is used as a search procedure. For non-solvable problems, the presence of Hamilton cycles in the associated digraphs means that coevolution will cycle indefinitely. For each solvable coevolutionary problem, there is a subset of solutions that dominate all other solutions in the solution space. In the context of games, coevolution would be able to search for these dominant strategies in the long run.

The second part of our framework has formalized coevolutionary processes as random walks on digraphs. These random walks are formulated as discrete time processes having the Markov property, which have allowed us to exploit the qualitative characterizations we have made on coevolutionary digraphs to construct relevant quantitative tools for analysis of those processes. We have shown that coevolution of solvable problems corresponds to an absorbing Markov chain with a single absorbing class. For this case, we have shown that the expected hitting time to the absorbing class is the quantity of interest. The other quantity is the limiting invariant distribution of the Markov chain. Although this is of principal interest for coevolution of non-solvable problems, the quantity can also be computed for the case of solvable problems as a result of Lemma 4.3.3.

We have demonstrated our quantitative tools through a combination of formal and computational case studies. We have derived expected hitting times and stationary distributions of specific coevolutionary Markov chains operating on transitive and regular tournaments. We have suggested an index for characterizing complexity in coevolutionary problems and show how a hierarchy of increasingly complex coevolutionary problems can be generated in a controlled manner that is motivated by statistical physics.

We believe this framework would provide the foundation for existing and the means to develop new tools for analysis and design of coevolutionary systems that use pairwise-comparison-based fitness measures. As a case in point, coevolution could *appear to be trapped* in cycles even when it is applied on solvable problems (as observed in various studies such as (Chong et al., 2012)). Although the associated reducible coevolutionary digraphs can be decomposed into acyclically ordered strong components (by Lemma 2.7.8), some of these strong components can be very large digraphs. In the extreme case, a reducible coevolutionary digraph of order $n$ has a strong component of $n - 1$ vertices. One possibility is that coevolution would cycle for a long time before *jumping* out and find the single dominant vertex. The other possibility is that the $(n - 1)$-sized component is the dominant one, in which case coevolution would seem to cycle indefinitely even though it has found (parts of) the dominant component. In the more realistic case, one could reframe the pairwise-comparison-based fitness measure in the context of *pareto dominance* and make the appropriate change to coevolution to

search in the direction of more dominant strong components (pareto fronts with higher true fitness values in the solution space) (de Jong and Pollack, 2004; de Jong, 2007).

The introduction of this framework for coevolutionary systems provides several directions for future studies. One of them is to provide a more definitive characterization of *problem hardness* affecting coevolutionary systems that include studies on local structures (internal cycle structures within strong components) of coevolutionary digraphs.

Beyond the restricted class of coevolutionary systems we have studied here, the mathematical methodologies we have introduced could be adopted in other formulations of coevolutionary systems. For example, those with population size greater than one and for which coevolutionary fitness measures other than pairwise-comparison (such as the use of average game payoffs) can be investigated. In this case, the entity of interest is the individual state (vertex) representing a finite collection (population) of pure strategies in the context of game-play. The advantage of such an approach is a more informative state space description. In contrast, in evolutionary game theoretic setting studied in the framework of continuous dynamical systems, one usually assumes infinite populations. However, in cases of complex, chaotic coevolutionary dynamics, this can be problematic (Ficici et al., 2005; Tiňo et al., 2013).

Our study is by no means the first one to consider representation of the structure of pairwise relations within the problem to be solved through a digraph. For example, in the multicomponent problems (Bonyadi et al., 2013; Klamroth et al., 2017), the digraph represents pairwise coupling between individual optimization problems that define together the global compound optimization problem to be solved. Typically, one can expect a small number of sub-problems (vertices) and there may not exist any coupling between a particular pair of sub-problems. In our case, the vertices represent strategies from a potentially huge set of strategies that can be chosen. In addition, the strategy set is fully coupled, i.e. between any two different strategies there exists an arc. Most importantly, direct application of Markov chains on such a digraph has the interpretation of a random walk on the strategy space following the outcomes of individual games.

## Acknowledgments

## References

Bader-Natal, A. and Pollack, J. B. (2005). Towards metrics and visualizations sensitive to coevolutionary failures. In *Coevolutionary and Coadaptive Systems Workshop: 2005 AAAI Symposium*, Washington DC, USA.

Bang-Jensen, J., Guo, Y., Gutin, G., and Volkmann, L. (1996). Paths, trees and cycles in tournaments. In Chartrand, G. and Jacobson, M., editors, *Surveys in Graph Theory*, Congressus Numerantium 115, pages 131–170. Springer-Verlag, San Francisco, CA.

Bang-Jensen, J., Guo, Y., Gutin, G., and Volkmann, L. (1997). A classification of locally semicomplete digraphs. *Discrete Mathematics*, 167-168:101–114.

Bang-Jensen, J. and Gutin, G. Z. (2009). *Digraphs Theory, Algorithms and Applications*. Springer-Verlag, London.

Bonyadi, M. R., Michalewicz, Z., and Barone, L. (2013). The travelling thief problem: The first step in the transition from theoretical problems to realistic problems. In

*Proc. IEEE 2013 Congress on Evolutionary Computation (CEC'13)*, pages 1037–1044, Cancun, Mexico.

Brualdi, R. A. (2010). Spectra of digraphs. *Linear Algebra and its Applications*, 432:2181–2213.

Cartlidge, J. and Bullock, S. (2004). Combating coevolutionary disengagement by reducing parasite virulence. *Evolutionary Computation*, 12(2):193–222.

Chong, S. Y., Tiňo, P., Ku, D. C., and Yao, X. (2012). Improving generalization performance in co-evolutionary learning. *IEEE Transactions on Evolutionary Computation*, 16(1):70–85.

Chung, K. L. (1960). *Markov Chains with Stationary Transition Probabilities*. Springer-Verlag, Berlin.

Cliff, D. and Miller, G. F. (1995). 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*, volume 929 of *Lecture Notes in Computer Science*, pages 200–218. Springer-Verlag.

Darwen, P. J. and Yao, X. (1997). Speciation as automatic categorical modularization. *IEEE Transactions on Evolutionary Computation*, 1(2):101–108.

de Jong, E. D. (2007). A monotonic archive for pareto-coevolution. *Evolutionary Computation*, 15(1):61–93.

de Jong, E. D. and Pollack, J. B. (2004). Ideal evaluation from coevolution. *Evolutionary Computation*, 12(2):159–192.

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

Floreano, D. and Nolfi, S. (1997). Adaptive behavior in competing co-evolving species. In Husbands, P. and Harvey, I., editors, *Proc. of the Fourth European Conference on Artificial Intelligence*, pages 378–387, Cambridge, MA. The MIT Press.

Grinstead, C. M. and Snell, J. L. (1997). *Introduction to Probability*. American Mathematical Society, Providence, Rhode Island.

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

Hemasinha, R. (2003). An algorithm to generate tournament score sequences. *Mathematical and Computer Modelling*, 37:377–382.

Hillis, W. D. (1990). Co-evolving parasites improve simulated evolution as an optimization procedure. *Physica D*, 42:228–234.

Iosifescu, M. (1980). *Finite Markov Processes and Their Applications*. John Wiley and Sons, New York.

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

Klamroth, K., Mostaghim, S., Naujoks, B., Poles, S., Purshouse, R., Rudolph, G., Ruzika, S., Sayin, S., Wiecek, M. M., and Yao, X. (2017). Multiobjective optimization for interwoven systems. *Journal of Multi-Criteria Decision Analysis*, 24(1-2):71–81.

Levin, D. A., Peres, Y., and Wilmer, E. L. (2009). *Markov Chains and Mixing Times*. American Mathematical Society, Providence, Rhode Island.

Meyer, C. D. (2000). *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia, PA.

Meyn, S. and Tweedie, R. L. (2009). *Markov Chains and Stochastic Stability*. Cambridge University Press, Cambridge, UK, 2 edition.

Moon, J. W. (1963). An extension of landau's theorem on tournaments. *Pacific Journal of Mathematics*, 4:1343–1345.

Moon, J. W. (1968). *Topics on Tournament*. Holt, Rinehart and Winston, New York.

Norris, J. R. (1998). *Markov Chains*. Cambridge University Press, Cambridge, UK.

Pirzada, S. (2012). Score structure in digraphs. In *An Introduction to Graph Theory*, pages 403–441. Orient Black Swan, Hyderabad, India.

Reid, K. B. and Zhang, C. Q. (1998). Score sequences of semicomplete digraphs. *Bulletin of the Institute of Combinatorics and its Applications*, 24:27–32.

Rosin, C. D. and Belew, R. K. (1997). New methods for competitive coevolution. *Evolutionary Computation*, 5(1):1–29.

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

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

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

Stanley, K. O. and Miikkulainen, R. (2002). 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)*, pages 242–248, New York, USA.

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