

Ranking graphs through hitting times of Markov chains

Emilio De Santis

Department of Mathematics, University of Rome La Sapienza, Rome, Italy

Correspondence

Emilio De Santis, Department of Mathematics, University of Rome La Sapienza, Piazzale Aldo Moro, 5, 00185 Rome, Italy.
Email: desantis@mat.uniroma1.it

Abstract

In the present paper we show that for any given digraph $\mathbb{G} = ([n], \vec{E})$, that is, an oriented graph without self-loops and 2-cycles, one can construct a 1-dependent Markov chain and n identically distributed hitting times T_1, \dots, T_n on this chain such that the probability of the event $T_i > T_j$, for any $i, j = 1, \dots, n$, is larger than $\frac{1}{2}$ if and only if $(i, j) \in \vec{E}$. This result is related to various *paradoxes* in probability theory, concerning in particular non-transitive dice.

KEYWORDS

1-dependent Markov chain, ordering, paradoxes in probability theory

1 | INTRODUCTION

Let us consider a collection of random variables $\mathcal{Y}_n = \{Y_1, \dots, Y_n\}$, defined on a same probability space and satisfying the no-tie condition, that is, any two of them are equal with probability zero. Let $[n] := \{1, 2, \dots, n\}$. For $i \neq j \in [n]$, one says that Y_i is *less than* (resp. is *equivalent to*) Y_j in the *stochastic precedence* sense if

$$\mathbb{P}(Y_i < Y_j) > \frac{1}{2}, \quad \left(\text{resp. } \mathbb{P}(Y_i < Y_j) = \frac{1}{2} \right).$$

This notion is natural in many applications, see for example [6, 8, 10] and references therein. From a theoretical point of view, it gives rise to a series of apparent paradoxes which are caused by the non-transitivity of the stochastic precedence comparison, see [11, 21, 23].

In order to present some results it is convenient to introduce the definition of *ranking graph* $\mathbb{G}(\mathcal{Y}_n) = ([n], \vec{E}(\mathcal{Y}_n))$ associated with the set of random variables $\mathcal{Y}_n = \{Y_1, \dots, Y_n\}$. For any $i, j \in [n]$, the pair (i, j) is an arrow of $\vec{E}(\mathcal{Y}_n)$ if and only if Y_j is less than Y_i in the stochastic precedence sense.

Although expressed in a different form, it is known that for any digraph $\mathbb{H} = ([n], \vec{E})$ there exists a set of n random variables \mathcal{Y}_n such that the ranking graph $\mathbb{G}(\mathcal{Y}_n)$ coincides with \mathbb{H} , that is, any set

of stochastic precedence relations can be achieved (see [19, 20]). This result was initially obtained in the field of voting theory by McGarvey, see [19], but it can be rewritten by using random variables and the stochastic precedence comparison between them (see [8, 14, 20]). We also mention [3, 4, 16, 22] which improve McGarvey's result by providing precise upper and lower bounds on the number of voters needed to achieve all possible ranking graphs.

In this paper we take a different point of view and we give a positive answer to the following problem: given a digraph $\mathbb{H} = ([n], \vec{E})$, can we find a stationary Markov chain with n *identically distributed* hitting times $\mathcal{T}_n = \{T_1, \dots, T_n\}$ such that $\mathbb{G}(\mathcal{T}_n) = \mathbb{H}$? Our interest in this problem is twofold.

- (1) The problem loses its combinatorial structure and becomes entirely probabilistic.
- (2) There are many applications where it is preferable to use hitting times rather than arbitrary random variables.

In various applications the type of random variables to use is constrained by the nature of the problem. For instance, in many competitive games the winner is the player who achieves his goal or target before the others. In this situation one is forced to work with hitting times. As an example of an application in this field, in the final part of the paper we define and analyze a Penney-type game in which these random variables arise quite naturally.

From a theoretical point of view, we emphasize that the structure of the hitting times is very particular and many properties on their distributions are known a priori, see for example [13, 18] and references therein. Moreover, the stationary Markov chains used in our construction have the particular property that the square of the transition matrix has all of its entries equal. Hence, they are *1-dependent uniform*, in the sense that they are 1-dependent [1, 2], with uniform invariant distribution.

The plan of the paper is as follows.

In Section 2 we give some basic notation and definitions, and preliminary results. In Section 3, we present our main result through an explicit construction of 1-dependent uniform Markov chains. In Section 4, we present a Penney-type games and we develop a qualitative analysis of it. In particular, we identify a threshold value beyond which it is possible to construct a Penney-type game unfavorable to the starting player whereas this is impossible below the same threshold.

2 | 1-DEPENDENT UNIFORM MARKOV CHAINS, PATTERNS AND IDENTICALLY DISTRIBUTED HITTING TIMES

First of all, recall that for a collection of r.v. $\mathcal{Y}_n = \{Y_1, \dots, Y_n\}$, we defined the *ranking graph* $\mathbb{G}(\mathcal{Y}_n) = ([n], \vec{E}(\mathcal{Y}_n))$ in the following way. For any $i, j \in [n]$,

$$(i, j) \in \vec{E}(\mathcal{Y}_n) \Leftrightarrow \mathbb{P}(Y_j < Y_i) > \frac{1}{2}. \quad (1)$$

It is clear that $\mathbb{G}(\mathcal{Y}_n)$ does not have loops or 2-cycles, thus it is a digraph.

We also recall the definition of 1-dependent uniform chain.

Definition 1. A Markov chain is a *1-dependent uniform chain* if it is stationary and the square of its transition matrix has all equal entries. The set of 1-dependent uniform chains is denoted by \mathcal{M}_1 .

We notice that a 1-dependent uniform chain $\mathbf{X} = (X_m : m \in \mathbb{N}_0)$ has the property that, for any sequence of increasing indices m_1, m_2, \dots with $m_{\ell+1} - m_{\ell} \geq 2$ (for $\ell \in \mathbb{N}$), the random variables

$(X_{m_\ell} : \ell \in \mathbb{N})$ are i.i.d. with X_{m_ℓ} distributed uniformly on the state space. In particular, the initial distribution is invariant and it coincides with the uniform distribution. For further properties on 1-dependent chains see [1,2], where such chains are studied in detail and characterized.

2.1 | Construction of 1-dependent uniform Markov chains.

Let $k, N \in \mathbb{N}$ with $N \geq 2$, we consider a sequence of i.i.d. random (column) vectors $\mathcal{U}_m = (U_{m,1}, \dots, U_{m,k})^T$, with $U_{m,i}$ i.i.d. uniformly distributed on $[N]$, for $i \in [k]$ and $m \in \mathbb{N}_0$. Next define the $k \times 2$ matrix

$$X_m^{(N,k)} = [\mathcal{U}_m, \mathcal{U}_{m+1}], \quad m \in \mathbb{N}_0. \quad (2)$$

Notice that the collection of random variables $\mathbf{X}^{(N,k)} = (X_m^{(N,k)} : m \in \mathbb{N}_0)$ forms a 1-dependent uniform chain, in particular it starts with the uniform distribution on the state space of all the matrices $k \times 2$ with elements belonging to $[N]$ (see Example 1). We notice that, for any $k \in \mathbb{N}$ and $N \geq 2$ the Markov chain $\mathbf{X}^{(N,k)}$ is not reversible, instead $\mathbb{P}(X_0^{(N,k)} = [\mathbf{1}, \mathbf{1}], X_1^{(N,k)} = [\mathbf{1}, \mathbf{2}]) = N^{-3k}$ but $\mathbb{P}(X_0^{(N,k)} = [\mathbf{1}, \mathbf{2}], X_1^{(N,k)} = [\mathbf{1}, \mathbf{1}]) = 0$, where $\mathbf{1}$ and $\mathbf{2}$ are column vectors with all the entries equal to 1 and 2, respectively.

2.2 | Patterns

For $M, k \in \mathbb{N}$, a *pattern* $Q = (q_{i,j} \in [M] \cup \{0\} : i \in [k], j \in [2])$ is a $k \times 2$ matrix, with the property that

$$q_{i,1} \in [M], \quad \text{for } i \in [k], \quad \text{and} \quad \sum_{i=1}^k \mathbf{1}_{\{q_{i,2} \neq 0\}} = 1.$$

For a pattern Q we define the *index of hump* as

$$h(Q) = j \quad \text{if } q_{j,2} > 0. \quad (3)$$

The collection of all the $k \times 2$ patterns with entries in $[M] \cup \{0\}$ is denoted by $\mathcal{P}_{M,k}$. In particular, any pattern in $\mathcal{P}_{M,k}$ has a number of entries different from zero which is equal to $k + 1$.

2.3 | Hitting time of a pattern

For any $h \in [k]$, we define the projection Π_h on the set of $k \times 2$ matrices in the following way: if $S = (s_{i,j} : i \in [k], j \in [2])$, then

$$\Pi_h(S) = \begin{bmatrix} s_{1,1} & 0 \\ \vdots & \vdots \\ s_{h-1,1} & 0 \\ s_{h,1} & s_{h,2} \\ s_{h+1,1} & 0 \\ \vdots & \vdots \\ s_{k,1} & 0 \end{bmatrix}. \quad (4)$$

Now let $\mathbf{X}^{(N,k)} = (X_m^{(N,k)} : m \in \mathbb{N}_0)$ be as in (2) and let $R \in \mathcal{P}_{M,k}$ be a target pattern, with $2 \leq M \leq N$. We define the hitting time of R as

$$T_R = \inf\{m \in \mathbb{N}_0 : \Pi_{h(R)}(X_m^{(N,k)}) = R\}, \quad (5)$$

in the following when $\Pi_{h(R)}(X_m^{(N,k)}) = R$ holds we will write $X_m^{(N,k)} \triangleright R$.

The hitting time T_R can be interpreted as the first time in which the pattern R occurs in the random sequence $(X_m^{(N,k)})_{m \in \mathbb{N}_0}$. From the finiteness of state space $[N]^k$ and the irreducibility of the Markov chain $\mathbf{X}^{(N,k)}$ it follows that T_R is finite almost surely.

2.4 | Overlap

For two different patterns

$$R = (r_{ij} \in [M] : i \in [k], j \in [2]), S = (s_{ij} \in [M] : i \in [k], j \in [2]) \in \mathcal{P}_{M,k}, \quad (6)$$

we define the *overlap* $O(R, S) \in \{0, 1\}^2$ in the following way

$$O_1(R, S) = \begin{cases} 1, & \text{if } r_{h,1} = s_{h,1}, \quad h \in [k] \text{ and } \delta_{h(R),h(S)}(r_{h(R),2} - s_{h(S),2}) = 0; \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

$$O_2(R, S) = \begin{cases} 1, & \text{if } r_{h(R),2} = s_{h(R),1}; \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

An analogous definition for strings has given in [17]. We notice that T_R and T_S are different w.p. 1 if and only if $O_1(R, S) = 0$. The meaning of $O_2(R, S)$ will be clarified by Lemmas 3 and 4 and Theorem 1 in which the relevance of $O_2(R, S)$ for the computation of $\mathbb{P}(T_R < T_S)$ becomes evident.

The overlap $O(R, S)$ is in general different from $O(S, R)$. Moreover it makes sense to consider the overlap of a pattern with itself, in which case the first component is always equal to 1. For patterns $R_1, \dots, R_n \in \mathcal{P}_{M,k}$, one has

$$\text{no-tie property of } T_{R_1}, \dots, T_{R_n} \Leftrightarrow \text{for distinct } i, j \in [n], \quad O_1(R_i, R_j) = 0. \quad (9)$$

In this case, for sake of simplicity, we say that the collection of patterns R_1, \dots, R_n is *no-tie*.

To familiarize with the notions and definitions we present the following example

Example 1. Let

$$R = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 2 & 0 \end{bmatrix}, \quad S = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 2 & 1 \end{bmatrix},$$

be two patterns in $\mathcal{P}_{2,3}$, with index of hump $h(R) = 2$ and $h(S) = 3$. The overlap $O(R, S) = (0, 1)$ and $O(S, R) = (0, 0)$. Thus the collection of patterns $\{R, S\}$ is no-tie.

Suppose that

$$\mathbf{X}^{(2,3)} = \begin{matrix} 1 & 1 & 1 & 1 & 1 & 2 & \dots \\ 2 & 2 & 1 & 1 & 2 & 1 & \dots \\ 2 & 1 & 2 & 1 & 2 & 1 & \dots \end{matrix}$$

Then $T_S = 2$ being

$$S \neq \Pi_3 \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 2 & 1 \end{bmatrix}, \quad S \neq \Pi_3 \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 1 & 2 \end{bmatrix}, \quad S = \Pi_3 \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 2 & 1 \end{bmatrix}.$$

Similarly, $T_R = 4$.

3 | RANKING GRAPHS THROUGH IDENTICALLY DISTRIBUTED HITTING TIMES

Let R be a pattern with $O(R, R) = (1, 0)$. Following the proof of Theorem 2.1 in [7] we recursively compute the discrete distribution of T_R .

Lemma 1. *Let k, M, N be integers such that $N \geq M \geq 2$ and $k \geq 1$. Let $R \in \mathcal{P}_{M,k}$ with $O(R, R) = (1, 0)$ and consider the 1-dependent uniform chain $\mathbf{X}^{(N,k)}$. Define $w(t) := \mathbb{P}(T_R = t)$, then the probabilities $(w(t) : t \in \mathbb{N}_0)$ are recursively determined from*

$$w(t) = N^{-k-1} - N^{-k-1} \sum_{s=0}^{t-2} w(s). \quad (10)$$

Proof. The event $\{X_t^{(N,k)} \triangleright R\}$ holds true with probability N^{-k-1} . Moreover, $\{X_t^{(N,k)} \triangleright R\}$ can be written as the union of the following three disjoint events:

- (i) $\{T_R = t\}$;
- (ii) $\{X_t^{(N,k)} \triangleright R\} \cap \{T_R = s\}$, for $s < t - 1$;
- (iii) $\{X_t^{(N,k)} \triangleright R\} \cap \{T_R = t - 1\}$.

For $s < t - 1$, as a consequence of 1-dependence, one has

$$\mathbb{P}(\{X_t^{(N,k)} \triangleright R\} \cap \{T_R = s\}) = w(s) \mathbb{P}(X_t^{(N,k)} \triangleright R) = N^{-k-1} w(s).$$

The event in (iii) has probability zero because, by hypothesis $O_2(R, R) = 0$. Therefore

$$N^{-k-1} = w(t) + N^{-k-1} \sum_{s=0}^{t-2} w(s),$$

which corresponds to (10). ■

Remark 1. Let $R_1, \dots, R_n \in \mathcal{P}_{M,k}$ such that $O(R_i, R_i) = (1, 0)$, for each $i \in [n]$. Let $N \geq M$ and consider the 1-dependent uniform chain $\mathbf{X}^{(N,k)}$. Then, by Lemma 1, the hitting times T_{R_1}, \dots, T_{R_n} are identically distributed because the distribution of any T_{R_i} is given by (10).

Given a digraph $\overline{\mathbb{G}} = ([n], \vec{E})$, we now construct a collection of associated patterns $\{R_u \in \mathcal{P}_{n+1, n+1} : u \in [n]\}$. For $\ell \in [n]$, pattern $R_\ell = (r_{ij}^{(\ell)} : i \in [n+1], j \in [2])$ is constructed in the following way:

1. $r_{1,1}^{(\ell)} = r_{\ell+1,2}^{(\ell)} = \ell$;
2. for any $j \in [n+1] \setminus \{\ell+1\}$, $r_{j,2}^{(\ell)} = 0$;
3. for any $j \in [n]$,

$$r_{j+1,1}^{(\ell)} = \begin{cases} j, & \text{if } (\ell, j) \in \vec{E}; \\ n+1, & \text{otherwise.} \end{cases} \quad (11)$$

We will say that the patterns R_1, \dots, R_n are *generated* by the graph $\overline{\mathbb{G}} = ([n], \vec{E})$.

First notice that $h(R_\ell) = \ell + 1$, for $\ell \in [n]$. In order to explain our construction of the patterns we observe that when $(i, j) \in \vec{E}$ then $O_2(R_i, R_j) = 0$ and $O_2(R_j, R_i) = 1$ (see Lemma 2), and this will cause $\mathbb{P}(T_{R_j} < T_{R_i}) > \frac{1}{2}$ through the phenomenon of clustering (see Lemmas 3 and 4, and Theorem 1). This phenomenon is analogous to what happens to the appearance of strings in a sequence of letters randomly drawn (see e.g., [9,17]). To illustrate the notation, we present the following example.

Example 2. Let us consider the graph $\overline{\mathbb{G}} = ([3], \vec{E})$ with $\vec{E} = \{(1, 3), (3, 2), (2, 1)\}$. Then the patterns R_1, R_2, R_3 generated by $\overline{\mathbb{G}}$ are

$$R_1 = \begin{bmatrix} 1 & 0 \\ 4 & 1 \\ 4 & 0 \\ 3 & 0 \end{bmatrix}, \quad R_2 = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 4 & 2 \\ 4 & 0 \end{bmatrix}, \quad R_3 = \begin{bmatrix} 3 & 0 \\ 4 & 0 \\ 2 & 0 \\ 4 & 3 \end{bmatrix}.$$

The present example will be continued at the end of the section where the patterns will be employed in the definition of three hitting times that show non-transitivity for the stochastic precedence.

We now consider the properties of the overlap for patterns R_1, \dots, R_n generated by a digraph $\overline{\mathbb{G}} = ([n], \vec{E})$.

Lemma 2. Let $n \geq 2$ and $\overline{\mathbb{G}} = ([n], \vec{E})$ be a digraph. The overlaps of the patterns R_1, \dots, R_n generated by $\overline{\mathbb{G}} = ([n], \vec{E})$ are

$$O(R_i, R_j) = \delta_{i,j} \cdot (1, 0) + (1 - \delta_{i,j}) \cdot \left[\mathbf{1}_{\{(i,j) \in \vec{E}\}} \cdot (0, 1) + \mathbf{1}_{\{(i,j) \notin \vec{E}\}} \cdot (0, 0) \right],$$

for any $i, j \in [n]$.

Proof. Case $i = j$, then $O_1(R_i, R_i) = 1$. The second component $O_2(R_i, R_i) = 0$ since $r_{i+1,1}^{(i)} = n + 1 \neq i = r_{i+1,2}^{(i)}$ (see (11)).

Case $i \neq j$. $O_1(R_j, R_i) = 0$ since $r_{1,1}^{(i)} = i \neq j = r_{1,1}^{(j)}$.

If $(i, j) \in \vec{E}$ then $O_2(R_j, R_i) = 1$. Indeed, the condition in (8) $r_{h(R^0),2}^{(j)} = r_{h(R^0),1}^{(i)}$ holds true since $r_{j+1,2}^{(j)} = r_{j+1,1}^{(i)} = j$.

If $(i, j) \notin \vec{E}$ then $O_2(R_j, R_i) = 0$. Indeed, the condition $r_{h(R^0),2}^{(j)} = r_{h(R^0),1}^{(i)}$ is not true since $r_{j+1,2}^{(j)} = j$ and $r_{j+1,1}^{(i)} = n + 1$. \blacksquare

Let us consider no-tie collection of patterns $\{R_1, \dots, R_\ell\}$ belonging to $\mathcal{P}_{M,k}$ and the corresponding hitting times $\mathcal{T}_\ell = \{T_{R_1}, \dots, T_{R_\ell}\}$ of $\mathbf{X}^{(N,k)}$, with $N \geq M$. We are interested to upper and lower bound

$$p_i(\mathcal{T}_\ell) := \mathbb{P} \left(\bigcap_{j \in [\ell]} \{T_{R_i} \leq T_{R_j}\} \right) \text{ for } i \in [\ell]. \tag{12}$$

First notice that by the no-tie property one has $\sum_{i \in [\ell]} p_i(\mathcal{T}_\ell) = 1$.

We also define the sequence of stopping times $(Z_h : h \in \mathbb{N}_0)$, as

$$Z_0 = \inf \{m \geq 0 : X_m^{(N,k)} \triangleright R_i \text{ for some } i \in [\ell]\} \tag{13}$$

and recursively, let

$$Z_{h+1} = \inf \{m \geq Z_h + 2 : X_m^{(N,k)} \triangleright R_i \text{ for some } i \in [\ell]\}. \tag{14}$$

It is immediate to notice that every hitting time Z_h is finite almost surely. We present this simple lemma without a proof.

Lemma 3. *For any $s \in \mathbb{N}$, one has*

$$\mathbb{P}(X_{Z_s+1}^{(N,k)} \triangleright R_i | X_{Z_s}^{(N,k)} \triangleright R_j) = \frac{O_2(R_j, R_i)}{N^k}. \tag{15}$$

The probabilities $(p_i(\mathcal{T}_\ell) : i \in [\ell])$ could be explicitly calculated through a linear system but for our purposes it will be more useful to have good upper and lower bounds.

Lemma 4. *Let $N \geq M \geq 2$ and $k, \ell \geq 2$. Let $R_1, \dots, R_\ell \in \mathcal{P}_{M,k}$ be a collection of no-tie patterns and let $O_2(R_i, R_i) = 0$, for $i \in [\ell]$. Let us take the uniform Markov chain $\mathbf{X}^{(N,k)}$ and the hitting times $\mathcal{T}_\ell = \{T_{R_i} : i \in [\ell]\}$. Then \mathcal{T}_ℓ are identically distributed. Moreover, for $i \in [\ell]$,*

$$p_i(\mathcal{T}_\ell) = \frac{v_i}{\sum_{j \in [\ell]} v_j}, \tag{16}$$

where

$$1 - \frac{1}{N^k} \sum_{j \in [\ell]} O_2(R_j, R_i) \leq v_i \leq 1 - \frac{1}{N^k} \left(1 - \frac{\ell - 1}{N^k} \right) \sum_{j \in [\ell]} O_2(R_j, R_i). \tag{17}$$

Proof. In the proof, we will write X_m for $X_m^{(N,k)}$. Lemma 1, Remark 1, and $O_2(R_i, R_i) = 0$, for each i , imply that \mathcal{T}_ℓ are identically distributed.

For any $i \in [\ell]$, by definition

$$p_i(\mathcal{T}_\ell) = \mathbb{P}(T_{R_i} = Z_0).$$

From the fact that the random variables $(X_m)_{m \in \mathbb{N}_0}$ are 1-dependent, we also have

$$p_i(\mathcal{T}_\ell) = \mathbb{P}(X_{Z_h} \triangleright R_i), \tag{18}$$

for any $h \in \mathbb{N}$. Moreover the times $(Z_s : s \in \mathbb{N}_0)$ are renewal times, that is,

$$\mathbb{P}(X_{Z_h} \triangleright R_i, Z_h - Z_{h-1} = s | X_{Z_{h-1}} = j, Z_{h-1} = t), \text{ for } h, s \in \mathbb{N}, \text{ and } i \in [\ell],$$

does not depend on $h, t \in \mathbb{N}$ and $j \in [n]$. This is again a consequence of the 1-dependent structure.

We define, for any $i \in [\ell]$, the sets of random times

$$\mathcal{V}_{i,t} := \{m < t : m = Z_s \text{ for some } s, X_m \triangleright R_i\}, \quad \mathcal{N}_{i,t} := \{s < t : X_s \triangleright R_i\}, \quad (19)$$

where $t \in \mathbb{N} \cup \{+\infty\}$. These sets' cardinalities are

$$V_{i,t} := |\mathcal{V}_{i,t}| = \sum_{s=0}^{\infty} \mathbf{1}_{\{Z_s \leq t-1\}} \mathbf{1}_{\{X_{Z_s} \triangleright R_i\}}, \quad N_{i,t} := |\mathcal{N}_{i,t}| = \sum_{s=0}^{t-1} \mathbf{1}_{\{X_s \triangleright R_i\}}. \quad (20)$$

By (18) and by the ergodic theorem for renewal process, one has

$$\lim_{t \rightarrow \infty} \frac{V_{i,t}}{\sum_{j \in [\ell]} V_{j,t}} = p_i(\mathcal{T}_\ell) \text{ a.s.}, \quad \lim_{t \rightarrow \infty} \frac{N_{i,t}}{t} = \frac{1}{N^{k+1}} \text{ a.s.} \quad (21)$$

We define the quantities $(v_i > 0 : i \in [\ell])$ as

$$v_i := \lim_{t \rightarrow \infty} \frac{V_{i,t} N^{k+1}}{t} \text{ a.s.}, \quad (22)$$

by hypothesis $N \geq M$ one has $v_i \leq 1$, for each $i \in [\ell]$. The equalities in (21) and the previous definition give (16).

We notice that

$$N_{i,t} = V_{i,t} + \sum_{j \in [\ell]: j \neq i} \sum_{s=0}^{\infty} \mathbf{1}_{\{Z_s \leq t-2\}} \mathbf{1}_{\{X_{Z_s} \triangleright R_j\}} \mathbf{1}_{\{X_{Z_s+1} \triangleright R_i\}}, \quad (23)$$

indeed if $X_{Z_s+1} \triangleright R_i$, for some s , the time $(Z_s + 1)$ belongs to $\mathcal{N}_{i,\infty}$ but it is not in $\mathcal{V}_{i,\infty}$. Let us multiply by N^{k+1}/t the previous formula and take the limit for $t \rightarrow \infty$, then, by the ergodic theorem, by (21) and (22), one obtains

$$1 = v_i + \lim_{t \rightarrow \infty} \frac{N^{k+1}}{t} \sum_{j \in [\ell]: j \neq i} \sum_{s=0}^{\infty} \mathbf{1}_{\{Z_s \leq t-2\}} \mathbf{1}_{\{X_{Z_s} \triangleright R_j\}} \mathbf{1}_{\{X_{Z_s+1} \triangleright R_i\}} \text{ a.s.} \quad (24)$$

By the ergodic theorem and Lemma 3 one has

$$1 = v_i + \sum_{j \in [\ell]: j \neq i} \frac{O_2(R_j, R_i)}{N^k} v_j. \quad (25)$$

Thus,

$$1 \leq v_i + \sum_{j \in [\ell]: j \neq i} \frac{O_2(R_j, R_i)}{N^k} = v_i + \sum_{j \in [\ell]} \frac{O_2(R_j, R_i)}{N^k}. \quad (26)$$

The inequality (26) corresponds to the first inequality in (17). In particular, $v_i \geq 1 - (\ell - 1)/N^k$, for any $i \in [\ell]$. Thus, for any fixed $i \in [\ell]$

$$\lim_{t \rightarrow \infty} \frac{N^{k+1}}{t} \sum_{s=0}^{\infty} \mathbf{1}_{\{Z_s \leq t-2\}} \mathbf{1}_{\{X_{Z_s} \triangleright R_i\}} \geq 1 - (\ell - 1)/N^k \quad a.s. \quad (27)$$

Now, by (23)–(27), Lemma 3 and ergodicity one has

$$v_i \leq 1 - \frac{1}{N^k} \left(1 - \frac{\ell - 1}{N^k} \right) \sum_{j \in [\ell]} O_2(R_j, R_i).$$

This end the proof. ■

We are now ready to present the following result on the construction of any digraph through the ranking graphs of 1-dependent uniform chains and identically distributed hitting times.

Theorem 1. *For any digraph $\overline{\mathbb{G}} = ([n], \vec{E})$ there exists $\mathbf{X} = (X_m : m \in \mathbb{N}_0) \in \mathcal{M}_1$ and a collection of identically distributed hitting times $\mathcal{T}_n = \{T_1, T_2, \dots, T_n\}$ on \mathbf{X} such that $\mathbb{G}(\mathcal{T}_n) = \overline{\mathbb{G}}$.*

Proof. Let us consider the 1-dependent uniform chain $\mathbf{X}^{(N, n+1)}$ with $N \geq n + 1$. By Lemma 2, the patterns R_1, \dots, R_n generated by $\overline{\mathbb{G}}$ have the no-tie property. Furthermore, by Lemmas 1 and 2 the hitting times $\{T_{R_1}, \dots, T_{R_n}\}$ are identically distributed.

For distinct indices i, j , one has

$$p_a(\{T_{R_i}, T_{R_j}\}) = \mathbb{P}(T_{R_a} = Z_1),$$

where $a \in \{i, j\}$.

In the case that (i, j) and (j, i) are not arrows of the digraph $\overline{\mathbb{G}}$ then, by Lemma 2, $O_2(R_i, R_i) = O_2(R_j, R_j) = O_2(R_i, R_j) = O_2(R_j, R_i) = 0$. Thus, by (23) of Lemma 4 follows that $N_{i,t} = V_{i,t}$ and $N_{j,t} = V_{j,t}$. Hence by (21) and (22) one has

$$p_i(\{T_{R_i}, T_{R_j}\}) = p_j(\{T_{R_i}, T_{R_j}\}) = \frac{1}{2}.$$

Hence, (i, j) and (j, i) does not belong to $\mathbb{G}(\mathcal{T}_n)$.

We now consider the case: (i, j) is in $\overline{\mathbb{G}}$. By Lemma 2, $O(R_i, R_j) = (0, 0)$ and $O(R_j, R_i) = (0, 1)$. Thus, by (23) follows that $N_{j,t} = V_{j,t}$ while

$$N_{i,t} = V_{i,t} + \sum_{s=0}^{\infty} \mathbf{1}_{\{Z_s \leq t-2\}} \mathbf{1}_{\{X_{Z_s} \triangleright R_j\}} \mathbf{1}_{\{X_{Z_s+1} \triangleright R_i\}}.$$

Now, defining v_i and v_j as in the proof of Lemma 4, $v_j = 1$ while

$$v_i \leq 1 - \frac{1}{N^k} \left(1 - \frac{1}{N^k} \right) < 1.$$

Therefore

$$p_i(\{T_{R_i}, T_{R_j}\}) = \frac{v_i}{v_i + v_j} < \frac{v_j}{v_i + v_j} = p_j(\{T_{R_i}, T_{R_j}\}).$$

Hence, (i, j) belongs to $\mathbb{G}(\mathcal{T}_n)$. ■

We end the section with the following example

Example 3. We want to construct $\mathbf{X} \in \mathcal{M}_1$ and three identically distributed hitting times such that

$$\mathbb{P}(T_1 < T_2) > \frac{1}{2}, \quad \mathbb{P}(T_2 < T_3) > \frac{1}{2}, \quad \mathbb{P}(T_3 < T_1) > \frac{1}{2}. \quad (28)$$

Thus we consider the digraph $\overline{\mathbb{G}}$ and the generated patterns R_1, R_2, R_3 defined in Example 2. We take $\mathbf{X} = \mathbf{X}^{(4,4)}$ and $T_i = T_{R_i}$, for $i = 1, 2, 3$. Now, by Theorem 1, the inequalities in (28) hold.

4 | A PENNEY-TYPE GAME

The classical Penney’s game concerns the occurrence of different strings in a sequence of independent random draws of letters. This kind of problem was studied and solved in [9, 17] (see also [12] for a version of the game with many players). In [17], among other results, the authors give the construction for the optimal reply or optimal string to every string chosen by the first player. The game is always unfavorable for the player who chooses first. The cause of this behavior lies in the absence of transitivity for the stochastic precedence order (see e.g., [10, 21, 23]).

To introduce our Penney-type game we need some notation. Let $\mathcal{T}_n = \{T_1, T_2, \dots, T_n\}$ be a collection of no-tie r.v., for $A \subset [n]$ we write

$$T^{(A)} = \min\{T_i : i \in A\}.$$

By the no-tie property, if the subsets $A, B \subset [n]$ are disjoint then $\mathbb{P}(T^{(A)} = T^{(B)}) = 0$.

Let $r_1, r_2 \in \mathbb{N}$, $\mathbf{X} \in \mathcal{M}_1$ and let \mathcal{T}_n be a collection of n identically distributed hitting times on \mathbf{X} with $n \geq r_1 + r_2$. We define the stochastic zero-sum game $G_{r_1, r_2}(\mathbf{X}, \mathcal{T}_n)$ as follows:

- Step 1.* Player I chooses a set $A \subset [n]$ with $|A|=r_1$.
- Step 2.* Player II chooses a set $B \subset [n] \setminus A$ with $|B|=r_2$.
- Step 3.* Player I chooses two nonempty sets $A' \subset A$ and $B' \subset B$.
- Step 4.* If $T^{(A')} < T^{(B')}$ then Player II pays $|B'|$ dollars to Player I, otherwise Player I pays $|A'|$ dollars to Player II.

The idea underlying this payoff is that, in the final stage, each player pays one dollar for betting on any hitting time and the winner takes all the stakes. After the choice of A' and B' , the expected payoff of Player I is given by

$$|B'| \cdot \mathbb{P}(T^{(A')} < T^{(B')}) - |A'| \cdot \mathbb{P}(T^{(A')} > T^{(B')}). \quad (29)$$

Note that for given $\mathbf{X} \in \mathcal{M}_1$ and for a collection of hitting times \mathcal{T}_n the expected payoff of the first player is a nondecreasing function of r_1 and r_2 , as long as $r_1 + r_2 \leq n$. Indeed, when $r'_1 \geq r_1$ or $r'_2 \geq r_2$, the first player can mimic, for $G_{r'_1, r'_2}(\mathbf{X}, \mathcal{T}_n)$, the strategies used in $G_{r_1, r_2}(\mathbf{X}, \mathcal{T}_n)$. Therefore his expected payoff is a monotone increasing function in r_1 and r_2 when the two players adopt an optimal strategy.

It is quite easy to construct for given r_1, r_2 , and $n \geq r_1 + r_2$ games of this kind that are fair or favorable to Player I. We will come back to this point in this section. However, we also determine a threshold $S(r_1, r_2)$ through a graph characterization that will be used in the following result (see formula (30)).

Theorem 2. For any $r_1, r_2 \in \mathbb{N}$, there exist $\mathbf{X} \in \mathcal{M}_1$ and a family of identically distributed hitting times \mathcal{T}_n on \mathbf{X} such that the game $G_{r_1, r_2}(\mathbf{X}, \mathcal{T}_n)$ is favorable to Player II if and only if $n \geq S(r_1, r_2)$.

This result is also related with “voting paradoxes,” see for example [3, 4].

4.1 | Existence of (r_1, r_2) -directional graphs

We start this subsection with some definitions. For a digraph $\mathbb{G} = ([n], \vec{E})$ and for disjoint $A, B \subset [n]$ we write $A \rightarrow B$ if for any $i \in A$ and $j \in B$ one has that $(i, j) \in \vec{E}$. For $r_1, r_2 \in \mathbb{N}$, we say that a digraph $\mathbb{G} = ([n], \vec{E})$ is (r_1, r_2) -directional if for any $A \subset [n]$ with $|A|=r_1$ there exists $B \subset [n] \setminus A$ with $|B|=r_2$ such that $A \rightarrow B$ (see [4, 15] for similar definitions). For any $r_1, r_2 \in \mathbb{N}$, let us define

$$S(r_1, r_2) := \inf \left\{ k \geq r_1 + r_2 : \text{there exists a } (r_1, r_2)\text{-directional tournament } ([k], \vec{E}) \right\}. \quad (30)$$

In [15] Erdős analyzes a problem that correspond to the existence of $(r_1, 1)$ -directional graphs (see also [5]). The probabilistic method developed there can be easily adapted in our case.

Theorem 3. For any $r_1, r_2 \in \mathbb{N}$ and any $n \geq S(r_1, r_2)$ there exists a (r_1, r_2) -directional tournament $\mathbb{T} = ([n], \vec{E})$. Moreover

$$S(r_1, r_2) \leq \inf \left\{ n \geq r_1 + r_2 : \binom{n}{r_1} \left(1 - \frac{1}{2^{r_1 r_2}} \right)^{\lfloor \frac{n-r_1}{r_2} \rfloor} < 1 \right\} < \infty. \quad (31)$$

Proof. For $r_1, r_2 \in \mathbb{N}$, we first assume that for a specific $n_0 \in \mathbb{N}$ there exists a (r_1, r_2) -directional tournament $\mathbb{T}_{n_0} = ([n_0], \vec{E})$. Then, we prove that for any $n > n_0$ there exists a (r_1, r_2) -directional tournament $\mathbb{T}_n = ([n], \vec{E}_n)$. The proof is by induction.

Suppose that for $n - 1 \geq n_0$ there is a (r_1, r_2) -directional tournament $\mathbb{T}_{n-1} = ([n - 1], \vec{E}_{n-1})$, then we will construct a tournament $\mathbb{T}_n = ([n], \vec{E}_n)$ that is (r_1, r_2) -directional.

For any distinct $i, j \in [n - 1]$ let (i, j) be in \vec{E}_n if and only if $(i, j) \in \vec{E}_{n-1}$. Moreover, for any $i \in [n - 1]$, we impose that (n, i) belongs to \vec{E}_n . It is clear that if \mathbb{T}_{n-1} is a (r_1, r_2) -directional tournament then also \mathbb{T}_n is an (r_1, r_2) -directional tournament. Indeed, if $A \subset [n - 1]$ with $|A|=r_1$ then one can select $B \subset [n - 1]$ with $|B|=r_2$ and $A \rightarrow B$, as in \mathbb{T}_{n-1} . On the other hand if we consider an $A \subset [n]$ such that $n \in A$ and $|A|=r_1$ then one can take $B \subset [n - 1]$ such that $(A \setminus \{n\}) \rightarrow B$ and $|B|=r_2$. In any case the relation $(A \setminus \{n\}) \rightarrow B$ implies $A \rightarrow B$ because $(n, i) \in \vec{E}_n$ for any $i \in [n - 1]$.

Now, we prove formula (31), by the probabilistic method (see e.g., [5]). For this purpose, we will construct a random tournament, denoted by $\mathbb{T}(n) = ([n], \vec{E}(n))$, and we will show that it is (r_1, r_2) -directional with positive probability.

For two distinct vertices u, v , either $(u, v) \in \vec{E}(n)$ or $(v, u) \in \vec{E}(n)$; both these events occur with probability 1/2. Moreover all the events involving distinct edges are assumed independent.

For given $r_1, r_2 \in \mathbb{N}$ let $\tilde{V} \subset [n]$ with $|\tilde{V}| = r_1$, we define the event

$$A_{\tilde{V}} := \{ \exists V' \subset [n] \setminus \tilde{V} : \tilde{V} \rightarrow V', \text{ with } |V'| = r_2 \}.$$

Now, for a given \tilde{V} having cardinality r_1 , let us choose a family of sets of vertices

$$\left(V_i : |V_i| = r_2, V_i \subset [n] \setminus \tilde{V}, \quad i = 1, \dots, \left\lfloor \frac{n - r_1}{r_2} \right\rfloor \right),$$

with $V_i \cap V_j = \emptyset$, for $i \neq j$.

By independence of the random directions involving different edges one has

$$\mathbb{P}(A_{\tilde{V}}^c) \leq \mathbb{P}\left(\bigcap_{i=1}^{\lfloor \frac{n-r_1}{r_2} \rfloor} \{\tilde{V} \rightarrow V_i\}^c\right) \leq \left(1 - \frac{1}{2^{r_1 r_2}}\right)^{\lfloor \frac{n-r_1}{r_2} \rfloor},$$

for any $\tilde{V} \subset [n]$ with $|\tilde{V}| = r_1$.

By subadditivity of the probability measure one has

$$\mathbb{P}\left(\bigcap_{\tilde{V} \subset [n]: |\tilde{V}|=r_1} A_{\tilde{V}}\right) = 1 - \mathbb{P}\left(\bigcup_{\tilde{V} \subset [n]: |\tilde{V}|=r_1} A_{\tilde{V}}^c\right) \geq 1 - \binom{n}{r_1} \left(1 - \frac{1}{2^{r_1 r_2}}\right)^{\lfloor \frac{n-r_1}{r_2} \rfloor}. \quad (32)$$

For any $r_1, r_2 \in \mathbb{N}$,

$$\lim_{n \rightarrow \infty} \binom{n}{r_1} \left(1 - \frac{1}{2^{r_1 r_2}}\right)^{\lfloor \frac{n-r_1}{r_2} \rfloor} = 0. \quad (33)$$

Formulas (32) and (33) imply (31). ■

Observe that, for a given $n \in \mathbb{N}$, if $\mathbb{G} = ([n], \vec{E})$ is a (r_1, r_2) -directional digraph and $\mathbb{G}' = ([n], \vec{E}')$ with $\vec{E} \subset \vec{E}'$ then also \mathbb{G}' is a (r_1, r_2) -directional digraph. This easy observation and Theorem 3 imply the following result.

Corollary 1. *For $r_1, r_2 \in \mathbb{N}$, all the (r_1, r_2) -directional digraphs have a number of vertices larger than or equal to $S(r_1, r_2)$.*

4.2 | Favorable, fair and unfavorable games

In the following, we take the point of view of the second player so we declare *favorable* (resp. *fair* and *unfavorable*) the game if the expected value of the payoff of Player II is positive (resp. null and negative), when both players adopt optimal strategies.

We first construct, for any $n \geq r_1 + r_2$ a fair game $G_{r_1, r_2}(\mathbf{X}, \mathcal{T}_n)$. Let $\mathbf{X} = (X_m : m \in \mathbb{N}_0)$ be a sequence of i.i.d. random variables taking value on $[n]$, with X_0 uniformly distributed on $[n]$, hence $\mathbf{X} \in \mathcal{M}_1$. Let $\mathcal{T}_n = \{T_1, \dots, T_n\}$ be the collection of identically distributed hitting times, where $T_i = \inf\{m \in \mathbb{N}_0 : X_m = i\}$, for $i \in [n]$. For any strategy of the players the game has a null expected payoff, therefore the game is trivially fair.

Now we construct an unfavorable game for any $n \geq r_1 + r_2$. Let us consider the Markov chain $\mathbf{X}^{(n+1, n+1)}$, a tournament $\mathbb{T} = ([n], \vec{E})$ with the property that $(i, n) \in \vec{E}$ for any $i \in [n-1]$, the patterns R_1, \dots, R_n generated by \mathbb{T} and the hitting times T_{R_1}, \dots, T_{R_n} . Player I takes A with $n \in A$ then, for any chosen set B by Player II, Player I selects $A' = \{n\}$ and $B' \subset B$ with $|B'|=1$. By construction $B \rightarrow \{n\}$, therefore Player I has guaranteed a positive expected payoff despite of the fact that this strategy could be suboptimal.

In order to find for which n there exist favorable games we need some more discussion and definitions. Definitions 2 and 3 are similar to others given in [11] but they are used there for different applications and purposes.

Definition 2. Let us consider two finite sets of random variables $S_A = \{S_i : i \in A\}$ and $S_B = \{S_i : i \in B\}$, such that $S_A \cup S_B$ has the no-tie property. We say that S_A is *small* with respect to S_B iff

$$\frac{1}{|A|} \sum_{i \in A} p_i(S_A \cup S_B) > \frac{1}{|B|} \sum_{i \in B} p_i(S_A \cup S_B). \tag{34}$$

First we present an example showing that the collective behavior cannot be deduced by pair relations.

Example 4. Let $S_1 = \frac{49}{100}$ and let S_2, S_3 be independent r.v. with uniform law on $[0, 1]$. The collection $\{S_1, S_2, S_3\}$ has the no-tie property. The r.v. S_1 is small with respect to S_i , for $i = 2, 3$, because $\mathbb{P}(S_1 < S_2) = \mathbb{P}(S_1 < S_3) = \frac{51}{100}$. But

$$\frac{1}{2} \sum_{i=2}^3 p_i(\{S_1, S_2, S_3\}) = \frac{1}{2} - \frac{1}{2} \left(\frac{51}{100}\right)^2 > \left(\frac{51}{100}\right)^2 = p_1(\{S_1, S_2, S_3\}).$$

Therefore $\{S_2, S_3\}$ is small with respect to S_1 . This example shows that the analysis of the smallness property cannot be reduced to the study of pair relations. In fact, these can be completely reversed when we move on to consider collections of random variables. A complete theory that studies all the possible ranking in a set of random variables is in [14, 20].

In order to avoid some difficulties in the construction of favorable games we define some special systems of random variables.

Definition 3. Let $S_n = \{S_1, \dots, S_n\}$ be a collection of random variables and let $\mathbb{G}(S_n) = ([n], \vec{E})$ be the associated ranking graph. We say that S_n is *2-determined* if for any two disjoint $A, B \subset [n]$ such that $A \rightarrow B$ then $S_B = \{S_i : i \in B\}$ is small with respect to $S_A = \{S_i : i \in A\}$.

Theorem 4. Let $n \geq 2$ and $\vec{\mathbb{G}} = ([n], \vec{E})$ be given. Let R_1, \dots, R_n be patterns in $\mathcal{P}_{n+1, n+1}$ generated by $\vec{\mathbb{G}} = ([n], \vec{E})$. For $N \geq n + 1$, let us consider the 1-dependent uniform chain $\mathbf{X}^{(N, n+1)}$ and the identically distributed hitting times $\mathcal{T}_n = \{T_{R_1}, \dots, T_{R_n}\}$. Then \mathcal{T}_n is 2-determined.

Proof. For any $C \subset [n]$ let $\mathcal{T}_C := \{T_i : i \in C\}$. For any disjoint $A, B \subset [n]$ with $A \rightarrow B$, we need to show

$$\frac{1}{|A|} \sum_{i \in A} p_i(\mathcal{T}_A \cup \mathcal{T}_B) < \frac{1}{|B|} \sum_{i \in B} p_i(\mathcal{T}_A \cup \mathcal{T}_B). \tag{35}$$

Inequality (35) holds true if and only if $\frac{1}{|B|} \sum_{i \in B} v_i - \frac{1}{|A|} \sum_{i \in A} v_i > 0$, for disjoint sets A, B .

For any $i \in A$, by Lemma 4, one has

$$v_i \leq 1 - \frac{1}{N^{n+1}} \left(1 - \frac{|A| + |B| - 1}{N^{n+1}}\right) \frac{|B|}{N^{n+1}}.$$

Hence,

$$\frac{1}{|A|} \sum_{i \in A} v_i \leq 1 - \frac{|B|}{N^{n+1}} + \frac{(|A| + |B| - 1)|B|}{N^{2n+2}}. \tag{36}$$

Analogously, by the first inequality in (17) and the hypothesis that $A \rightarrow B$, one has

$$\frac{1}{|B|} \sum_{i \in B} v_i \geq 1 - \frac{1}{N^{n+1}} \sum_{i \in B} \sum_{j \in B \setminus \{i\}: (i,j) \in \vec{E}} 1. \quad (37)$$

Hence,

$$\frac{1}{|B|} \sum_{i \in B} v_i \geq 1 - \frac{|B| - 1}{2N^{n+1}} \quad (38)$$

Therefore one obtains

$$\frac{1}{|B|} \sum_{i \in B} v_i - \frac{1}{|A|} \sum_{i \in A} v_i > \frac{|B|}{2N^{n+1}} - \frac{(|A| + |B| - 1)|B|}{N^{2n+2}} > |B| \cdot \left[\frac{1}{2N^{n+1}} - \frac{1}{N^{2n+1}} \right], \quad (39)$$

where the last inequality is a consequence of $N > |A| + |B| - 1$. Thus, for any $N \geq n + 1 \geq 3$ and for any choice of nonempty disjoint sets $A, B \subset [n]$ the l.h.s. in (39) results larger than zero. ■

We give the key result that will allow the construction of favorable games.

Theorem 5. *Let $r_1, r_2 \in \mathbb{N}$, $n \geq r_1 + r_2$, let $\mathbf{X} \in \mathcal{M}_1$ and $\mathcal{T}_n = \{T_1, \dots, T_n\}$ be a collection of identically distributed hitting times.*

i. The game $G_{r_1, r_2}(\mathbf{X}, \mathcal{T}_n)$ is favorable $\Rightarrow \mathbb{G}(\mathcal{T}_n)$ is (r_1, r_2) -directional.

Moreover, let us suppose that \mathcal{T}_n is 2-determined then

ii. The game $G_{r_1, r_2}(\mathbf{X}, \mathcal{T}_n)$ is favorable $\Leftrightarrow \mathbb{G}(\mathcal{T}_n)$ is (r_1, r_2) -directional.

Proof. Item i. The proof is by contradiction. Suppose that $\mathbb{G}(\mathcal{T}_n) = ([n], \vec{E}(\mathcal{T}_n))$ is not (r_1, r_2) -directional. By hypothesis, Player I can select A , with cardinality r_1 , such that, for any B with cardinality r_2 , $A \not\leftrightarrow B$ in $\mathbb{G}(\mathcal{T}_n)$. Let us consider such A and B . Now, Player I selects $i \in A$ and $j \in B$ such that $(i, j) \notin \vec{E}(\mathcal{T}_n)$. Hence, $\mathbb{P}(T_j < T_i) \leq \frac{1}{2}$. Then, by choosing $A' = \{i\}$ and $B' = \{j\}$ Player I has guaranteed that the game is either fair or in his favor. Obviously, this strategy could be suboptimal for Player I. In any case, the game $G_{r_1, r_2}(\mathbf{X}, \mathcal{T}_n)$ is not favorable for Player II.

Item ii. For any set A having cardinality r_1 , Player II selects B , with cardinality r_2 , such that $A \rightarrow B$. Let us consider such a B . From the fact that the system is 2-determined for any choice of no empty $A' \subset A$ and $B' \subset B$ one obtains

$$\frac{1}{|A'|} \sum_{i \in A'} p_i(\mathcal{T}_{A' \cup B'}) < \frac{1}{|B'|} \sum_{i \in B'} p_i(\mathcal{T}_{A' \cup B'}).$$

By (29), the expected payoff of Player II is positive, for all $A' \subset A$, $B' \subset B$. ■

Now we prove Theorem 2 presented in the beginning of this section. For given r_1, r_2 if $n < S(r_1, r_2)$, by Corollary 1 any graph is not (r_1, r_2) -directional; then by Theorem 5, the considered game is fair or unfavorable.

If $n \geq S(r_1, r_2)$, by Theorem 3, one takes a digraph $\overline{\mathbb{G}} = ([n], \vec{E})$ that is (r_1, r_2) -directional. Let us take a chain $\mathbf{X}^{(N, n+1)} \in \mathcal{M}_1$ with $N \geq n + 1$. Then the patterns $R_1, \dots, R_n \in \mathcal{P}_{n+1, n+1}$ generated by $\overline{\mathbb{G}}$ are considered. By Theorem 1, we know that $\mathbb{G}(\{T_{R_1}, \dots, T_{R_n}\}) = \overline{\mathbb{G}}$, moreover T_{R_1}, \dots, T_{R_n} are

identically distributed hitting times and, by Theorem 4, these hitting times are 2-determined. Finally, by the item ii of Theorem 5, the game $G_{r_1, r_2}(\mathbf{X}^{(N, n+1)}, \{T_{R_1}, \dots, T_{R_n}\})$ is favorable. This proves Theorem 2.

ACKNOWLEDGMENTS

The author would like to thank the anonymous reviewers for several helpful comments and suggestions that have improved the text. The research was partially supported by Fondi di Ateneo 2018.

REFERENCES

1. J. Aaronson, D. Gilat, and M. Keane, *On the structure of 1-dependent Markov chains*, J. Theoret. Probab. **5** (1992), 545–561.
2. J. Aaronson, D. Gilat, M. Keane, and V. de Valk, *An algebraic construction of a class of one-dependent processes*, Ann. Probab. **17** (1989), 128–143.
3. N. Alon, *Voting paradoxes and digraphs realizations*, Adv. Appl. Math. **29** (2002), 126–135.
4. N. Alon, G. Brightwell, H. A. Kierstead, A. V. Kostochka, and P. Winkler, *Dominating sets in k-majority tournaments*, J. Combin. Theory Ser. B **96** (2006), 374–387.
5. N. Alon and J. H. Spencer, *The probabilistic method*, in *Wiley-Interscience Series in Discrete Mathematics and Optimization*, 2nd ed., Wiley-Interscience [John Wiley & Sons], New York, NY, 2000.
6. M. A. Arcones and F. J. Samaniego, *On the asymptotic distribution theory of a class of consistent estimators of a distribution satisfying a uniform stochastic ordering constraint*, Ann. Statist. **28** (2000), 116–150.
7. G. Blom and D. Thorburn, *How many random digits are required until given sequences are obtained?* J. Appl. Probab. **19** (1982), 518–531.
8. C. R. Blyth, *Some probability paradoxes in choice from among random alternatives*, J. Amer. Statist. Assoc. **67** (1972), 366–373.
9. R. Chen and A. Zame, *On fair coin-tossing games*, J. Multivariate Anal. **9** (1979), 150–156.
10. E. De Santis, F. Fantozzi, and F. Spizzichino, *Relations between stochastic orderings and generalized stochastic precedence*, Probab. Eng. Inform. Sci. **29** (2015), 329–343.
11. E. De Santis, Y. Malinovsky, and F. Spizzichino, *Stochastic precedence and minima among dependent variables*, Methodol. Comput. Appl. Probab. (2020), 1–19. <https://doi.org/10.1007/s11009-020-09772-3>
12. E. De Santis and F. Spizzichino, *First occurrence of a word among the elements of a finite dictionary in random sequences of letters*, Electron. J. Probab. **17** (2012), 9.
13. E. De Santis and F. Spizzichino, *Usual and stochastic tail orders between hitting times for two Markov chains*, Appl. Stoch. Models Bus. Ind. **32** (2016), 526–538.
14. E. De Santis and F. Spizzichino, *Load-sharing dependence models and construction of voting situations for any arbitrary ranking schemes*, arXiv:2004.09417 [math.PR], 2020, pp. 1–27.
15. P. Erdős, *On a problem in graph theory*, Math. Gaz. **47** (1963), 220–223.
16. P. Erdős and L. Moser, *On the representation of directed graphs as unions of orderings*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **9** (1964), 125–132.
17. L. J. Guibas and A. M. Odlyzko, *String overlaps, pattern matching, and nontransitive games*, J. Combin. Theory Ser. A **30** (1981), 183–208.
18. F. Manzo and E. Scoppola, *Exact results on the First hitting via conditional strong quasi-stationary times and applications to metastability*, J. Stat. Phys. **174** (2019), 1239–1262.
19. D. C. McGarvey, *A theorem on the construction of voting paradoxes*, Econometrica **21** (1953), 608–610.
20. D. G. Saari, *A chaotic exploration of aggregation paradoxes*, SIAM Rev. **37** (1995), 37–52.
21. R. P. Savage Jr., *The paradox of nontransitive dice*, Amer. Math. Mon. **101** (1994), 429–436.
22. R. Stearns, *The voting problem*, Amer. Math. Mon. **66** (1959), 761–763.
23. S. Trybuła, *Cyclic random inequalities*, Appl. Math. **10** (1969), 123–127.

How to cite this article: De Santis E. Ranking graphs through hitting times of Markov chains. *Random Struct Alg* 2021;1–15. <https://doi.org/10.1002/rsa.20998>