HAL open science

The frequency of pattern occurrence in random walks

Sergi Elizalde, Megan Martinez

To cite this version:

Sergi Elizalde, Megan Martinez. The frequency of pattern occurrence in random walks. 27th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2015), Jul 2015, Daejeon, South Korea. pp.217-228, 10.46298/dmtcs.2476 . hal-01337824

HAL Id: hal-01337824

https://hal.science/hal-01337824v1

Submitted on 27 Jun 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

FPSAC 2015, Daejeon, South Korea DMTCS proc. **FPSAC'15**, 2015, 217–228

The frequency of pattern occurrence in random walks

Sergi Elizalde[1†] and Megan Martinez[1‡]

[1] *Department of Mathematics, Dartmouth College, Hanover, NH 03755.*

Abstract. In the past decade, the use of ordinal patterns in the analysis of time series and dynamical systems has become an important tool. Ordinal patterns (otherwise known as a permutation patterns) are found in time series by taking n data points at evenly-spaced time intervals and mapping them to a length-n permutation determined by relative ordering. The frequency with which certain patterns occur is a useful statistic for such series. However, the behavior of the frequency of pattern occurrence is unstudied for most models. We look at the frequency of pattern occurrence in random walks in discrete time, and we define a natural equivalence relation on permutations under which equivalent patterns appear with equal frequency, regardless of probability distribution. We characterize these equivalence classes applying combinatorial methods.

Résumé. Au cours de la dernière décennie, l'utilisation des motifs ordinaux dans l'analyse des séries chronologiques et systèmes dynamiques est devenu un outil important. Des motifs ordinaux (autrement appelés motifs de permutations) se trouvent dans les séries chronologiques en prenant n points de données au intervalles de temps uniformément espacées et les faisant correspondre à une permutation de longueur n déterminée par leur ordre relatif. La fréquence avec laquelle certains motifs apparaissent est une statistique utile pour ces séries. Toutefois, le comportement de la fréquence d'apparition de ces motifs n'a pas été étudié pour la plupart des modèles. Nous regardons la fréquence d'occurrence des motifs dans les marches aléatoires en temps discret, et nous définissons une relation d'équivalence naturelle sur des permutations dans laquelle les motifs équivalents apparaissent avec la même fréquence, quelle que soit la distribution de probabilité. Nous caractérisons ces classes d'équivalence utilisant des méthodes combinatoires.

Keywords: permutation pattern; random walk; time series analysis; ordinal pattern; pattern frequency

1 Introduction

Time series analysis deals with the extraction of information from sequences of data points, typically measured at uniform time intervals. Understanding the characteristics of the data enables better predictions of the future behavior of a phenomenon. There are a number of different statistical methods that can be applied to the study of time series. A relatively new method involves the analysis of its "ordinal patterns." This approach, pioneered in the dynamical systems community by Bandt, Keller and Pompe [5, 4] and

†Email: `sergi.elizalde@dartmouth.edu`. Partially supported by grant #280575 from the Simons Foundation and by grant H98230-14-1-0125 from the NSA.

‡Email: `megan.a.martinez.gr@dartmouth.edu`.

1365–8050 © 2015 Discrete Mathematics and Theoretical Computer Science (DMTCS), Nancy, France

surveyed by Amigó [1], is particularly amenable to a combinatorial treatment. For one-dimensional deterministic time series that arise from iterating a map, combinatorial analyses of the ordinal patterns for specific maps have appeared in [2, 7, 8, 3]. However, perhaps surprisingly, very little is known about the behavior of ordinal patterns in a random setting.

In this paper we study, from a combinatorial perspective, the ordinal patterns that occur in random walks. We define an equivalence relation on permutations, which we conjecture determines when two patterns occur with the same probability in any random walk, and we provide a combinatorial characterization of the equivalence classes. We expect such a characterization to be useful in many of the applications of ordinal patterns in random walks that have recently appeared in the dynamical systems literature. Such applications include the analysis of stock indices and economic indicators, both to quantify the randomness of certain time periods in the series [10], and to show that the degree of market inefficiency is correlated with the number of missing patterns [11]. In the related setting of Gaussian processes with stationary increments, the frequency of ordinal patterns has been estimated in [9] and computed exactly for some small patterns in [6].

Permutation patterns are found in a time series by taking n data points at evenly-spaced time intervals and mapping them to a length-n permutation determined by relative ordering. For example, a sequence $4.8, -4.1, 3.1, 5.2$ would map to the permutation $3124 \in S_4$. The frequencies of the patterns that occur are measured and used to make conclusions about the behavior of the data. Central to this analysis is an understanding of the frequency with which patterns occur in a random time series.

Among the different models that are used for random time series, one of the most basic and applicable is a one-dimensional random walk in discrete time. To construct such a walk, take $n-1$ independent and identically distributed (i.i.d.) continuous random variables $X_1, X_2, \ldots, X_{n-1}$; we call these *steps*. The walker starts at 0, and at time i the walker is at $X_1+X_2+\ldots+X_i$. It is easy to see that not all permutation patterns occur with equal probability in such a random walk. For example, if the X_i's are chosen from a distribution that only takes positive values, then the pattern $123\ldots n$ will occur with probability 1.

Define a map $p : \mathbb{R}^{n-1} \to S_n$ where $p(X_1, \ldots, X_{n-1}) = \pi$ if the entries of the permutation π have the same relative ordering as the walk $Z_0, Z_1, \ldots, Z_{n-1}$, where $Z_i = X_1 + \cdots + X_i$ for all i (with $Z_0 = 0$). More precisely, $p(X_1, \ldots, X_{n-1}) = \pi$ if $\pi(i) = |\{k \mid Z_k \leq Z_i\}|$ for all i. If the associated random walk contains repeated values, i.e. $Z_i = Z_j$ for some i, j, it will be our convention to leave p undefined. Since we only deal with continuous distributions, the probability of having repeated values is zero.

While not all permutations occur in the image of p with equal probability, it turns out that there are certain classes of permutations that do have equal probability of occurring, regardless of the probability distribution chosen for the X_i's. For example, the pattern 132 will always occur with equal probability as 213, since $p(X_1, X_2) = 132$ if and only if $p(X_2, X_1) = 213$. In general, the reverse-complement of a permutation will occur with equal probability as the permutation itself. It turns out that such equivalencies are not restricted to reverse-complements, but are littered across S_n. For example, 1432 and 2143 also occur with equal probability because $p(X_1, X_2, X_3) = 1432$ if and only if $p(X_2, X_1, X_3) = 2143$.

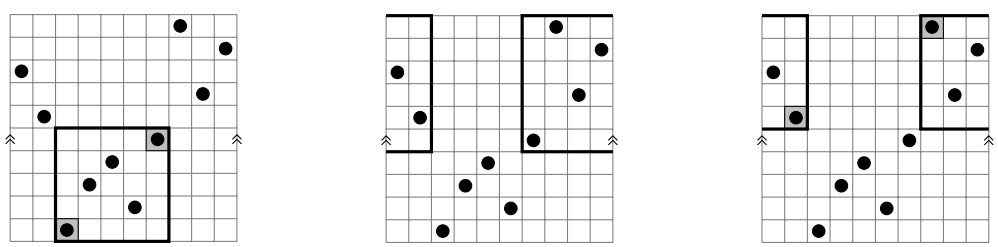
In this paper we define an equivalence relation on permutations that arises naturally from the above examples. It has the property that any two permutations π and τ in the same equivalence class satisfy $\mathbb{P}(p(X_1, \ldots, X_{n-1}) = \pi) = \mathbb{P}(p(X_1, \ldots, X_{n-1}) = \tau)$ for any continuous probability distribution on the X_i (as long as they are i.i.d.). Our main result, stated in Section 2, gives a complete characterization of the equivalence classes. Sections 3, 4, and 5 introduce the tools and ideas used in the proof. Finally, in Section 6 we conjecture that the converse property holds, that is, if two permutations occur equally often for any continuous probability distribution on the X_i, then they are equivalent under our definition.

2 Equivalence on Permutations

We define a natural equivalence relation on permutations $\pi, \tau \in S_n$, which is suggested by the above examples. We let $\pi \sim \tau$ if there exists some $\rho \in S_{n-1}$ such that for every $x_1, \ldots, x_n \in \mathbb{R}$, we have $p(x_1, \ldots, x_{n-1}) = \pi$ if and only if $p(x_{\rho(1)}, \ldots, x_{\rho(n-1)}) = \tau$. If $X_1, \ldots, X_{n-1}$ are i.i.d. random variables, then the sequences $(X_1, \ldots, X_{n-1})$ and $(X_{\rho(1)}, \ldots, X_{\rho(n-1)})$ have the same joint probability distribution, and so $\pi \sim \tau$ implies that $\mathbb{P}(p(X_1, \ldots, X_{n-1}) = \pi) = \mathbb{P}(p(X_1, \ldots, X_{n-1}) = \tau)$ for any continuous probability distribution on the random variables X_i.

The main result of this paper precisely characterizes the equivalence classes for $\sim$. Our characterization is best illustrated by displaying permutations $\pi \in S_n$ on an $n \times n$ grid by filling the boxes $(i, \pi(i))$ for $1 \leq i \leq n$ with a dot. Our convention is that the (i, j) box is in the ith column from the left and the jth row from the bottom, as in cartesian coordinates.

Let us introduce a few definitions. A *block* in a permutation is a set of consecutive entries whose values also form a consecutive set. On the grid, a block is a square subgrid with a dot in each row and column, which implies that the regions right, left, above, and below a block are empty. A *cylindrical block* is a generalization of this notion where the requirement of consecutive positions is relaxed by considering 1 and n to be consecutive. If we identify the left and right edges of our $n \times n$ grid, then a cylindrical block is a square subgrid of the resulting cylinder with a dot in each row and column. Note that a cylindrical block can be either a regular block or a block that spans the left and right sides of the grid. We say that a cylindrical block is *bordered* if it has at least three entries, and the entries with highest and lowest value occur precisely at the outer positions, regardless of which one is on which side (see Figure 1).



(a) The bordered cylindrical block $1\,3\,4\,2\,5$.

(b) The unbordered cylindrical block $5\,10\,7\,9\,8\,6$.

(c) The bordered cylindrical block $10\,7\,9\,8\,6$.

Fig. 1: Examples of bordered and unbordered cylindrical blocks in $\pi = 8\,6\,1\,3\,4\,2\,5\,10\,7\,9$.

Given a permutation that contains a bordered cylindrical block, we can generate another permutation by performing a *flip* on the bordered cylindrical block. For a regular block, a flip is simply a 180° rotation of the contents of the block. For a cylindrical block that spans the left/right boundary, a flip is akin to a 180° rotation, except the entries on the left side of the block are rotated and moved to the right and vice-versa, while the entries that are not part of the block are shifted right or left accordingly (see Figure 2). Recall that the reverse-complement operation on π corresponds to a 180° rotation of the whole $n \times n$ grid, which is a block, but not necessarily bordered. Our main theorem characterizes equivalence classes in S_n in terms of flips:

Theorem 2.1 *Let $\pi, \tau \in S_n$. Then $\pi \sim \tau$ if and only if τ can be obtained from π through a sequence of flips of bordered cylindrical blocks and the reverse-complement operation.*

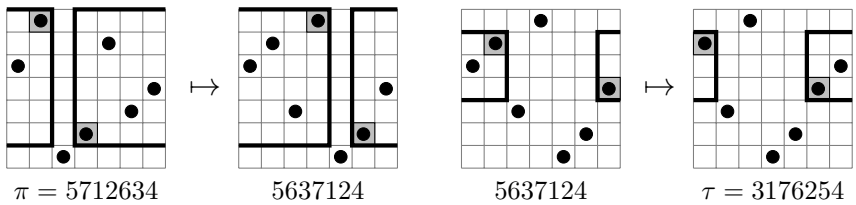
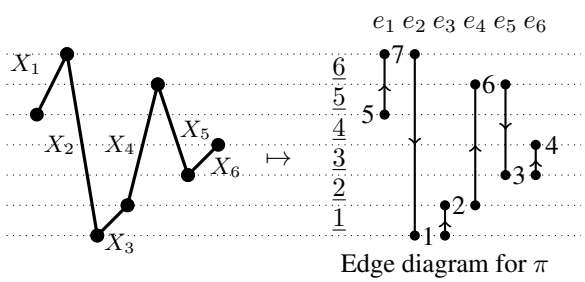


Fig. 2: A sequence of flips of bordered cylindrical blocks that maps $\pi = 5712634$ to $\tau = 6371245$.

Proving this result requires us to delve into the structure of permutations. Our aim is to understand how we can permute the steps of one permutation (seen as a random walk) to get another. For real numbers $x_1, \ldots, x_{n-1}$ to satisfy $p(x_1, \ldots, x_{n-1}) = \pi$, there are some forced relationships among them. For example, if $p(x_1, x_2, x_3) = 1423$, then $x_1 + x_2 > 0$ and $x_1 > x_3$, among other relations.

It will be convenient to draw each step of π (seen as a walk) as a vertical, directed edge. In the coordinate plane, we draw edge e_i as the line segment that connects $(i, \pi(i))$ to $(i, \pi(i+1))$, directed upwards if $\pi(i+1) > \pi(i)$ and downwards otherwise. Define $\mathrm{sgn}(e_i) = \mathrm{sgn}(\pi(i+1) - \pi(i))$. We denote the (unordered) set of $n-1$ edges corresponding to π by $\mathcal{E}_\pi$ and call it its *edge diagram* (see Figure 3). We think of the y-coordinates $1, 2, \ldots, n$ as vertices, and of e_i as a directed edge from $\pi(i)$ to $\pi(i+1)$. With this interpretation, a sequence of edges $e_i, e_{i+1}, \ldots, e_{j-1}$ forms a path from vertex $\pi(i)$ to vertex $\pi(j)$.



Edge diagram for π

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

$L(\pi)$

Fig. 3: The permutation $\pi = 5712634$, its edge diagram, and its matrix $L(\pi)$.

We partition the y-axis of our edge diagram into the intervals $\underline{1} = [1, 2]$, $\underline{2} = [2, 3]$, $\ldots$, $\underline{n-1} = [n-1, n]$, that we call *levels*. We can then write each edge as a formal sum of the levels it covers: $e_i = \sum_{j=\pi(i)}^{\pi(i+1)-1} \underline{j}$ when $\pi(i) < \pi(i+1)$ and $e_i = -\sum_{j=\pi(i+1)}^{\pi(i)-1} \underline{j}$ when $\pi(i+1) < \pi(i)$. Notice that letting $L(\pi)$ be the $(n-1) \times (n-1)$ matrix with entries $-1, 0, 1$ where

$$(L(\pi))_{ij} = \begin{cases} \mathrm{sgn}(\pi(i+1) - \pi(i))), & \text{if } \pi(i) \le j < \pi(i+1) \text{ or } \pi(i+1) \le j < \pi(i), \\ 0, & \text{otherwise,} \end{cases}$$

then we have that $e_i = \sum_{j=1}^{n-1} (L(\pi))_{ij} \cdot \underline{j}$, that is,

$$\begin{bmatrix} e_1 \\ \vdots \\ e_{n-1} \end{bmatrix} = L(\pi) \cdot \begin{bmatrix} \underline{1} \\ \vdots \\ \underline{n-1} \end{bmatrix}. \tag{1}$$

For ease of notation, we write $\underline{j} \in e_i$ if $(L(\pi))_{ij} \neq 0$.

The following lemma, whose proof is omitted in this extended abstract, states that L is a group homomorphism between S_n and $GL_{n-1}(\mathbb{C})$. Thus, L gives a representation of the symmetric group, which can be shown to be isomorphic to the standard representation.

Lemma 2.2 *For every $\pi, \tau \in S_n$, we have $L(\tau\pi) = L(\pi)L(\tau)$.*

Noticing that L maps the identity permutation to the identity matrix, it follows from Lemma 2.2 that $L(\pi^{-1}) = L(\pi)^{-1}$, and, in particular, that $L(\pi)$ is always invertible. Additionally, since $L(\pi)$ and $L(\pi^{-1})$ are matrices with integral entries, they have integral determinants, and so $\det(L(\pi)) = \pm 1$.

In order to classify equivalence classes, we first show that $\pi \sim \tau$ if and only if $L(\pi)$ and $L(\tau)$ are related by permutations of rows and columns. In the following lemma, P_ρ denotes the permutation matrix associated to ρ; that is, $(P_\rho)_{ij}$ equals 1 if $\rho(i) = j$ and 0 otherwise.

Lemma 2.3 *For $\pi, \tau \in S_n$, $\pi \sim \tau$ if and only if there exist $\sigma, \rho \in S_{n-1}$ such that $P_{\rho^{-1}} L(\pi) P_\sigma = L(\tau)$.*

The proof of this lemma uses the fact that if we let $D_\pi = \{\mathbf{x} \in \mathbb{R}^{n-1} \mid p(\mathbf{x}) = \pi\}$, then multiplication by $L(\pi)$ is a bijection from $\mathbb{R}^{n-1}_{>0}$ to D_π. Thus, if $\rho \in S_{n-1}$ is such that $(x_1, \ldots, x_{n-1}) \in D_\pi$ if and only if $(x_{\rho(1)}, \ldots, x_{\rho(n-1)}) \in D_\tau$, then multiplication by $L(\pi^{-1})P_\rho L(\tau)$ is a bijection from $\mathbb{R}^{n-1}_{>0}$ to itself. Since this matrix and its inverse have non-negative integer entries, it must be a permutation matrix P_σ.

If we consider the results of Lemma 2.3 in the context of an edge diagram, σ is applied to the levels and ρ is applied to the edges. Since the edges of an edge diagram are unordered, we only need to consider permutations of the levels by σ, as will be stated in Lemma 2.4.

We need a notation that describes edges in terms of levels. For vertices $s, t \in [n]$ with $s < t$ in an edge diagram, we write the *interval* between s and t as $[s, t]$ and define this to be the union of levels $\underline{s} \cup \underline{s+1} \cup \ldots \cup \underline{t-1}$. We write an edge $e_i \in \mathcal{E}_\pi$ as a directed interval $[\pi(i), \pi(i+1)]_+$ (if $\pi(i) < \pi(i+1)$) or $[\pi(i+1), \pi(i)]_-$ (if $\pi(i) > \pi(i+1)$) for $1 \leq i \leq n-1$; the subscripts indicate upwards or downwards direction, respectively. For example, in Figure 3, the edges in $\mathcal{E}_\pi$ are $e_1 = [5, 7]_+, e_2 = [1, 7]_-, e_3 = [1, 2]_+, e_4 = [2, 6]_+, e_5 = [6, 3]_-, e_6 = [3, 4]_+$. Given an edge $e = [i, j]_\pm \in \mathcal{E}_\pi$ and any interval $[s, t]$, we write $[s, t] \subseteq e$ if $[s, t] \subseteq [i, j]$ and $e \subseteq [s, t]$ if $[i, j] \subseteq [s, t]$.

Permuting the levels of an edge diagram by σ takes level $\underline{i}$ and moves it to height $\underline{\sigma(i)}$. Intervals and edges are shifted accordingly: the interval $[i, j]$ is moved to $\sigma.[i, j] = \underline{\sigma(i)} \cup \underline{\sigma(i+1)} \cup \ldots \cup \underline{\sigma(j-1)}$ (note that $\sigma.[i, j]$ may no longer be an interval), and the edge $[i, j]_\pm$ is moved to $(\sigma.[i, j])_\pm$, where the sign is preserved. The set of images of elements of $\mathcal{E}_\pi$ is denoted by $\mathcal{E}_{\sigma.\pi}$, and we say that this is a *well-defined* edge diagram if for every edge $[i, j]_\pm \in \mathcal{E}_\pi$, $\sigma.[i, j]$ is an interval. If $\mathcal{E}_{\sigma.\pi}$ is the edge diagram of a permutation, we say that $\mathcal{E}_{\sigma.\pi}$ is a *proper* edge diagram and call the corresponding permutation $\sigma.\pi$.

Recall that the edge diagram of a permutation τ forms a path $\tau(1), \tau(2), \ldots, \tau(n)$ where the vertices $\tau(i)$ and $\tau(i+1)$ are connected by an edge e_i. A well-defined edge diagram $\mathcal{E}_{\sigma.\pi}$ is proper if and only if its edges form a path. This allows us to ignore the order of the edges in $\mathcal{E}_\pi$, since π can be recovered from the path. This idea is useful when proving the following statement:

Lemma 2.4 *Given $\pi, \tau \in S_n$, $\pi \sim \tau$ if and only if there exists $\sigma \in S_{n-1}$ such that $\mathcal{E}_{\sigma.\pi} = \mathcal{E}_\tau$.*

Example 2.5 *Let $\pi = 54621873, \tau = 73218465 \in S_8$, whose edge diagrams are drawn in Figure 4, and let $\sigma = 2365471 \in S_7$. Then $\mathcal{E}_\pi = \{[4, 5]_-, [4, 6]_+, [2, 6]_-, [1, 2]_-, [1, 8]_+, [7, 8]_-, [3, 7]_-\}$ and $\mathcal{E}_{\sigma.\pi} = \{[3, 7]_-, [2, 3]_-, [1, 2]_-, [1, 8]_+, [4, 8]_-, [4, 6]_+, [5, 6]_-\} = \mathcal{E}_\tau$. Therefore $\pi \sim \tau$.*

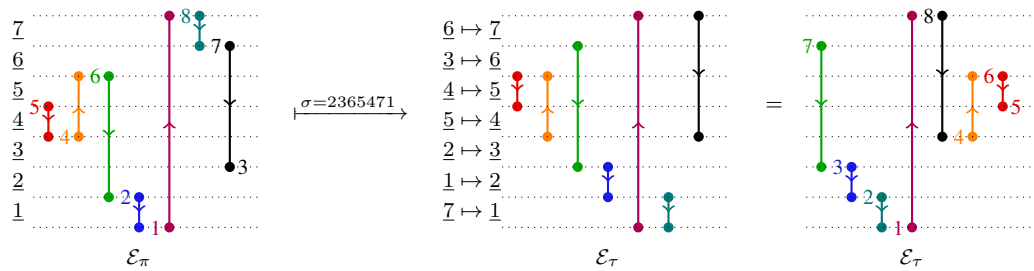


Fig. 4: $\pi = 54621873$ is mapped to $\tau = 73218463$ by applying $\sigma = 2365471$. The edges of $\mathcal{E}_\tau$ are first displayed in the same order as in $\mathcal{E}_\pi$, and then rearranged to form a path.

Definition 2.6 *Let $\pi \in S_n$ and $\sigma \in S_{n-1}$ such that $\mathcal{E}_{\sigma.\pi}$ is well-defined. We say that σ* is valid with respect to π *if $\mathcal{E}_{\sigma.\pi}$ is proper. If σ is valid with respect to π, we will say that σ acts validly on π.*

One of the goals of the next few sections to describe all σ that are valid with respect to π and understand how they transform π. Note that if $\sigma_1, \sigma_2 \in S_{n-1}$ are such that $\mathcal{E}_{\sigma_1.\pi}$ and $\mathcal{E}_{\sigma_2.(\sigma_1.\pi)}$ are proper edge diagrams, then $\mathcal{E}_{\sigma_2.(\sigma_1.\pi)} = \mathcal{E}_{(\sigma_2\sigma_1).\pi}$. Allowing edge diagrams that are not well-defined (meaning the edges are not contiguous) would give a group action of S_{n-1} on S_n, but for our purposes it is convenient to restrict only to well-defined diagrams.

3 Valid Flips

In this section, we define an operation on edge diagrams that is analogous to a flip of a bordered cylindrical block. The remaining sections will then focus on proving that any two equivalent permutations differ by a sequence of these operations.

We will use the following two properties of edge diagrams:

(1) Let $\{p_\pi, q_\pi\} = \{\pi(1), \pi(n)\}$ be the endpoints in the edge diagram for π (assume $p_\pi < q_\pi$). Then for any $i \in [n-1]$, the number of edges that contain $\underline{i}$ is odd if $\underline{i} \in [p_\pi, q_\pi]$ and even otherwise. So, if σ is valid with respect to π, then $\sigma.[p_\pi, q_\pi]$ must be an interval.

(2) Define a cycle in an edge diagram to be a sequence of vertices $v_1, v_2, v_3, \ldots, v_k$ where either $[v_i, v_{i+1}]_+$ or $[v_{i+1}, v_i]_-$ is an edge for every i, and either $[v_k, v_1]_+$ or $[v_1, v_k]_-$ is an edge. If $\mathcal{E}_{\sigma.\pi}$ is well-defined, then it does not contain a cycle.

Property (1) is straightforward. To prove property (2), one can show that a cycle in $\mathcal{E}_{\sigma.\pi}$ would require the existence of a cycle in $\mathcal{E}_\pi$.

Property (1) suggests a slight alteration to our edge diagrams. Since $[p_\pi, q_\pi]$ must remain an interval, any σ that is valid with respect to π must treat $[p_\pi, q_\pi]$ as if it were a "phantom" edge. Let $[p_\pi, q_\pi]_\pm$ be an edge such that $\mathrm{sgn}([p_\pi, q_\pi]_\pm) = 1$ if $\pi(n) < \pi(1)$ and $\mathrm{sgn}([p_\pi, q_\pi]_\pm) = -1$ if $\pi(1) < \pi(n)$. Define a *modified edge diagram* on a permutation π to be $\mathcal{E}^\star_\pi = \mathcal{E}_\pi \cup [p_\pi, q_\pi]^\star_\pm$. The ornamentation on $[p_\pi, q_\pi]^\star_\pm$ is needed to ensure there is a bijective correspondence between permutations and modified edge diagrams. This added edge, which converts the path of $\mathcal{E}_\pi$ into a cycle, will be represented with a dashed line in a drawing of $\mathcal{E}^\star_\pi$ (see Figure 5). However, $[p_\pi, q_\pi]^\star_\pm$ is mathematically treated like any other edge.

The terminology from edge diagrams can be naturally extended to modified edge diagrams. For instance, we say that $\sigma.\mathcal{E}_\pi^\star = \{\sigma.e \mid e \in \mathcal{E}_\pi^\star\}$ is well-defined if $\sigma.[i,j]$ is an interval for every edge $[i,j]_\pm \in \mathcal{E}_\pi^\star$. Additionally, $\sigma.\mathcal{E}_\pi^\star$ is *proper* if $\mathcal{E}_{\sigma.\pi}$ is proper.

Lemma 3.1 *Let $\pi \in S_n$ and $\sigma \in S_{n-1}$ such that $\sigma.\mathcal{E}_\pi^\star$ is a well-defined edge diagram. Then $\sigma.\mathcal{E}_\pi^\star$ forms a cycle if and only if $\mathcal{E}_{\sigma.\pi}$ is a proper edge diagram.*

The proof of Lemma 3.1 requires using property (2) to show that $\mathcal{E}_{\sigma.\pi}$ will form a path whenever $\sigma.\mathcal{E}_\pi^\star$ forms a cycle. As a consequence of Lemma 3.1, working with modified edge diagrams is analogous to working with edge diagrams. Additionally, $\sigma.\mathcal{E}_\pi^\star$ will be notated as $\mathcal{E}_{\sigma.\pi}^\star$.

Define a *flip* to be a permutation $\sigma \in S_{n-1}$ of the form

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & \cdots & n-1 \\ 1 & 2 & 3 & \cdots & i-1 & j-1 & j-2 & \cdots & i & j & \cdots & n-1 \end{pmatrix}$$

where $j > i+1$. We say that σ *flips* the interval $[i,j]$.

Since adjacent transpositions are a particular case of flips, it is clear that flips generate S_{n-1}. Given any interval $[i,j]$ in $\mathcal{E}_\pi^\star$, we say that $[i,j]$ *respects* $\mathcal{E}_\pi^\star$ if for every $e \in \mathcal{E}_\pi^\star$, $e \subseteq [i,j]$, $[i,j] \subseteq e$, or $[i,j] \cap e = \emptyset$.

Lemma 3.2 *Suppose that $\sigma \in S_{n-1}$ is a flip of the interval $[i,j]$ (where $j > i+1$). Then σ is valid with respect to π if and only if $[i,j]$ respects $\mathcal{E}_\pi^\star$.*

We omit the proof of this lemma. The idea is that the cycle in $\mathcal{E}_\pi^\star$ behaves well in relation to the interval $[i,j]$. Therefore, flipping $[i,j]$ will simply reorder certain portions of the cycle.

The intervals mentioned in Lemma 3.2 correspond to bordered cylindrical blocks. We therefore call a flip as described in Lemma 3.2 a *valid flip* (with respect to π) and the interval it flips a *valid interval* (in $\mathcal{E}_\pi^\star$). In general, we say that σ transforms π by a *sequence of valid flips* if there exist $\sigma_1, \ldots, \sigma_l \in S_{n-1}$ with $\sigma_l\sigma_{l-1}\cdots\sigma_1 = \sigma$ such that σ_i is a valid flip with respect to $\sigma_{i-1}\sigma_{i-2}\ldots\sigma_1.\pi$ for every i.

Lemma 3.3 *For every $\pi \in S_n$, the interval $[i,j]$ is valid in $\mathcal{E}_\pi^\star$ if and only if in the grid of π, the values $[i,j]$ (which are in positions $\pi^{-1}(i), \pi^{-1}(i+1), \ldots, \pi^{-1}(j)$) form a bordered cylindrical block or $[i,j] = [1,n]$. In this case, flipping the valid interval $[i,j]$ in $\mathcal{E}_\pi^\star$ is equivalent to flipping the corresponding bordered cylindrical block in the grid of π, or, when $[i,j] = [1,n]$, to taking the reverse-complement.*

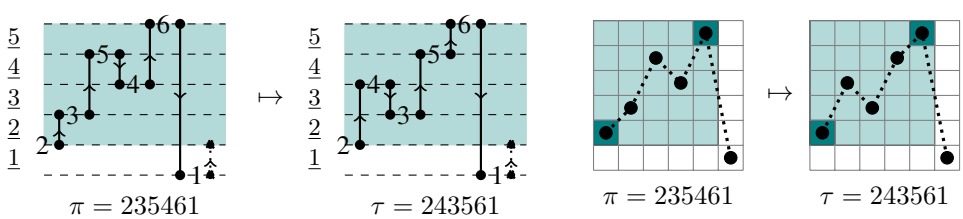


Fig. 5: Flipping valid interval $[2,6]$ in the edge diagram for $\pi = 235461$ yields the same result as flipping the bordered cylindrical block with values $[2,6]$ (and positions $[1,5]$) in the grid of π.

For an example of Lemma 3.3, see Figure 5. Now we can restate Theorem 2.1 as follows:

Theorem 3.4 *Let $\pi, \tau \in S_n$. Then $\pi \sim \tau$ if and only if τ can be obtained from π by a sequence of valid flips.*

In order to prove this theorem, we need to further explore the structure of edge diagrams. The aim of the definitions and lemmas in the following section will be to decompose our edge diagrams into nested structures to which we will be able to apply an inductive argument.

4 Irreducible Intervals

For given $\pi \in S_n$, most permutations of the levels of its edge diagram are not valid. Our induction argument will rely on the ability to partition the levels of the edge diagram into intervals that remain intervals under any valid permutation of the levels. We first introduce the idea of an irreducible interval, which is a maximal interval whose levels remain adjacent under the action of any valid permutation. Then we show that every edge diagram can be uniquely partitioned into intervals of this type.

Definition 4.1 *Let $\pi \in S_n$. An interval $[s,t]$ in the modified edge diagram of π is said to satisfy the* linked conditions *if, for every $\sigma \in S_{n-1}$ that is valid with respect to π, the following are true:*

(1) The image $\sigma.[s,t]$ is an interval.

(2) If we let $i,j \in [n]$ be such that $\sigma.[s,t] = [i,j]$, then the ordered tuple $(\sigma(s), \sigma(s+1), \sigma(s+2), \ldots, \sigma(t-1))$ equals either $(i, i+1, \ldots, j-1)$ or $(j-1, j-2, \ldots, i)$.

Additionally, $[s,t]$ is called an irreducible interval *if it is a maximal interval that satisfies the linked conditions (i.e. for any interval $[x,y]$ such that $[s,t] \subsetneq [x,y]$, $[x,y]$ does not satisfy the linked conditions).*

Note that any proper subinterval of an irreducible interval will satisfy the linked conditions but will fail the maximality condition.

Lemma 4.2 *For any $\pi \in S_n$, the interval $[1,n]$ in the modified edge diagram of π can be uniquely partitioned into irreducible intervals.*

Proof: Suppose that $[s,t]$ and $[u,v]$ are irreducible intervals with a nonempty intersection. By the maximality condition, $[s,t]$ cannot properly contain $[u,v]$ and vice versa. Without loss of generality, assume that $s \leq u$ and $t \leq v$. Then $[s,t] \cap [u,v] = [u,t]$. Since both $[s,t]$ and $[u,v]$ satisfy the linked conditions, every σ that is valid with respect to π maps $[s,t]$ and $[u,v]$ to intervals. Since these intervals have nonempty intersection, it follows that σ maps $[s,t] \cup [u,v] = [s,v]$ to an interval. In fact, the second linked condition for $[s,t]$ and $[u,v]$ quickly implies that $[s,v]$ satisfies the condition as well. By the maximality of irreducible intervals, this is only possible if $s = u$ and $t = v$. Therefore, any two irreducible intervals are either disjoint or equal.

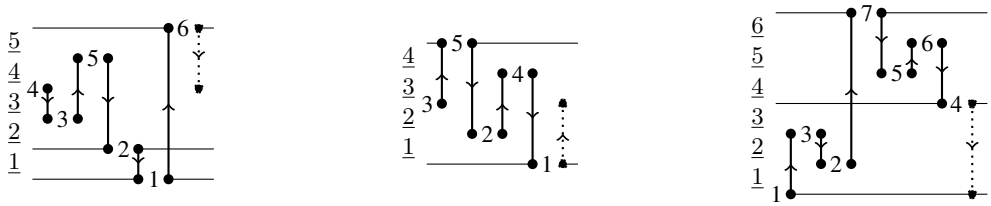
Since intervals of width 1 trivially satisfy the linked conditions, it is clear that every level of the edge diagram is contained in some irreducible interval. The condition on maximality for irreducible intervals implies that the partition is unique. □

We let $\mathcal{I}_\pi = \{[x_i, x_{i+1}] : 0 \leq i < k\}$, where $x_0 = 1$ and $x_k = n$, denote the partition of the levels of some $\mathcal{E}^\star_\pi$ into irreducible intervals. We call $\mathcal{I}_\pi$ the *irreducible partition* of π, and we call the x_i its borders.

It is not true that every irreducible interval is valid; for instance, interval $[1,4]$ in Figure 6(c) is irreducible but not valid. However, any valid interval is the union of adjacent irreducible intervals. Indeed, given an interval $[i,j]$ that is valid (so $j > i+1$) with respect to $\mathcal{E}^\star_\pi$, we know by Lemma 3.2 that if σ flips $[i,j]$, then $\mathcal{E}^\star_{\sigma.\pi}$ is proper. When $j \neq i+1$, this implies that i and j are borders of the irreducible partition $\mathcal{I}_\pi$, since $\sigma.(\underline{i-1} \cup \underline{i}) = \underline{i-1} \cup \underline{j-1}$ and $\sigma.(\underline{j-1} \cup \underline{j}) = \underline{i} \cup \underline{j}$ are not intervals.

Finding irreducible intervals using Definition 4.1 is impractical. A corollary to Theorem 3.4, omitted in this extended abstract, states that irreducible intervals are completely determined by the valid intervals. More precisely, any interval $[a, b]$ satisfies the linked conditions if and only if $[a, b]$ is contained in or disjoint from every valid interval in $\mathcal{E}^\star_\pi$. Thus, x_i is a border of $\mathcal{I}_\pi$ if and only if x_i is an endpoint of some valid interval. The proof is not obvious; a correspondence between valid intervals in equivalent permutations is established in order to derive these facts.

Although a few permutations only have irreducible intervals of width one, such as $1\,2\,3 \ldots n$ and $1\,n\,(n-1) \ldots 2$, permutations often have wide irreducible intervals. For instance, whenever an edge diagram has an edge of width two, the two levels covered by this edge must remain adjacent, thereby creating an irreducible interval with width at least 2. In fact, many permutations have one single irreducible interval $[1, n]$. Examples of irreducible intervals are given in Figure 6.



(a) The modified edge diagram for 435216 has two irreducible intervals: $[1, 2]$ and $[2, 6]$.

(b) The modified edge diagram for 35241 has one irreducible interval: $[1, 5]$.

(c) The modified edge diagram for 1327564 has two irreducible intervals: $[1, 4]$ and $[4, 7]$.

Fig. 6: Examples of irreducible intervals

Definition 4.1 and Lemma 4.2 allow us to consider signed permutations of the irreducible intervals of π, rather than unsigned permutations of the levels, since every σ that is valid with respect to π permutes and possibly flips its irreducible blocks. Recall that a signed permutation is a bijection $\mu : [-n] \cup [n] \to [-n] \cup [n]$ (where $[-n] \cup [n] = \{-n, \ldots, -1, 1, \ldots, n\}$) such that $\mu(-i) = -\mu(i)$ for all $i \in [n]$. We write the signed permutation μ as $\mu(1)\mu(2) \ldots \mu(n)$, leaving out the images of the negative numbers since they follow, and writing $\overline{\mu(i)}$ instead of $-\mu(i)$. Note that the barring operation is an involution, i.e. $\overline{\overline{\mu(i)}} = \mu(i)$. Let $|\mu(i)|$ denote the entry $\mu(i)$ without a bar. We denote the set of signed permutations of length k by B_k, and the infinite set of all signed permutations by B.

If $\pi \in S_n$ has k irreducible intervals, we can apply a signed permutation $\mu \in B_k$ to the edge diagram of π as follows. Each entry of μ describes where the corresponding irreducible interval of π is moved, and a barred entry indicates that the order of the levels inside the irreducible interval is reversed. As an example, see Figure 7. We denote the image of applying μ to the irreducible intervals of a modified edge diagram, $\mathcal{E}^\star_\pi$, by $\mathcal{E}^\star_{\mu.\pi}$. If, additionally, this is a proper modified edge diagram, we say that μ is *valid* with respect to π and denote the corresponding permutation by $\mu.\pi$.

We need to generalize our notion of a valid flip to signed permutations. If $[x_i, x_j]$ is a valid interval in $\mathcal{E}^\star_\pi$, then we say that $\mu = 1\,2 \cdots i\,\overline{j}\,\overline{j-1} \cdots \overline{i+1}\,j{+}1 \cdots k \in B_k$ is a *valid flip* with respect to π. We say that μ transforms π by a *sequence of valid flips* if there exist $\mu_1, \ldots, \mu_l \in B_k$ with $\mu_l \mu_{l-1} \cdots \mu_1 = \mu$ such that μ_i is a valid flip with respect to $\mu_{i-1}\mu_{i-2} \ldots \mu_1.\pi$ for all i.

We can now restate our main theorem using the irreducible partition.

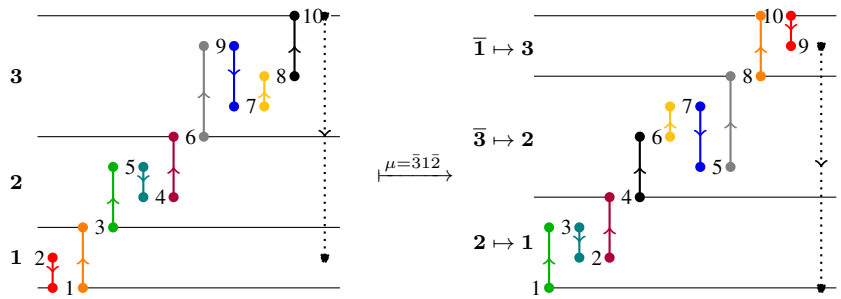


Fig. 7: 2 1 3 5 4 6 9 7 8 10 is mapped to 1 3 2 4 7 5 6 8 10 9 by applying $\mu = \bar{3}1\bar{2}$.

Theorem 4.3 *Let $\pi, \tau \in S_n$ where π has k irreducible intervals. Then $\pi \sim \tau$ if and only if there exists some $\mu \in B_k$ such that $\mathcal{E}^\star_{\mu.\pi} = \mathcal{E}^\star_\tau$ and μ transforms π by a sequence of valid flips.*

5 Cohesive Intervals and Partitions

In this section, we introduce the notion of a cohesive partition, a generalization of the irreducible partition. We are going to induct on the number of blocks of a partition of the edge diagram of π, so the purpose of this generalization is to have a method of coarsening the irreducible partition.

Definition 5.1 *Let $\pi \in S_n$ with k irreducible intervals. We say that an interval $[a, b]$ is* cohesive *in π if (1) $[a, b]$ is a valid interval in $\mathcal{E}^\star_\pi$, and (2) for every $\mu \in B_k$ that is valid with respect to π, $\mu.[a, b]$ is an interval.*

Condition (2) is difficult to check, since one would in principle have to verify the property for all signed permutations of length k. Notice that every edge in $\mathcal{E}^\star_\pi$ satisfies condition (2), but most will fail (1). All irreducible intervals satisfy condition (2), but not necessarily condition (1), as is the case for interval $[1, 4]$ in Figure 6(c). Some examples of cohesive intervals are given in Figure 8; notice that $[1, 7]$ and $[6, 10]$ are valid intervals but not cohesive. Since all cohesive intervals are valid, they are the union of consecutive irreducible intervals; we will sometimes use the borders x_i of the irreducible partition to describe them. For convenience, we write $\{a_0, a_1, \ldots, a_l\}_<$ to denote the partition into intervals $[a_i, a_{i+1}]$ for $0 \leq i < l$.

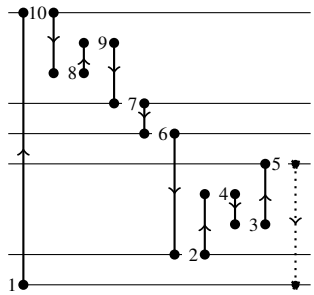


Fig. 8: In $\mathcal{E}^\star_\pi$, where $\pi = 1\,10\,8\,9\,7\,6\,2\,4\,3\,5$, the intervals $[1, 6]$, $[1, 10]$, $[2, 5]$, and $[7, 10]$ are cohesive.

Definition 5.2 *A partition $\mathcal{P} = \{x_s = a_0, a_1, a_2, \ldots, a_l = x_t\}_<$ of a cohesive interval $[x_s, x_t]$ is called a* cohesive partition *if $[a_i, a_{i+1}]$ is irreducible or cohesive for every $1 \leq i < l$.*

Cohesive partitions provide us a more general setting to prove our main theorem: we will consider signed permutations of the blocks in $\mathcal{P}$ that result in a valid modified edge diagram, and show that all such permutations are sequences of valid flips. We can then prove Theorem 4.3 as a special case, using that $[1, n]$ is a cohesive interval and the irreducible partition is a cohesive partition.

In the rest of this section, we assume that $\pi \in S_n$ has k irreducible intervals, and we use $\mathcal{P} = \{a_0, a_1, a_2, \ldots, a_l\}_<$ to denote a cohesive partition of the cohesive interval $[x_s, x_t]$. The notation $[x_s, x_t]_\mathcal{P}$ is used when considering $[x_s, x_t]$ as a union of the blocks of $\mathcal{P}$. This is in contrast to $[x_s, x_t]$, which is thought of as a union of levels.

In order to consider signed permutations of the blocks for a partition $\mathcal{P}$ in a cohesive interval $[x_s, x_t]$, we need to introduce a few definitions. Given $\mu \in B_k$, its *reverse-complement* is the signed permutation with $\mu^{RC}(i) = \overline{\mu(k-i+1)}$ for $1 \leq i \leq k$. For $\alpha_1, \alpha_2, \ldots, \alpha_k \in B$, we define the *inflation* of μ by $\alpha_1, \alpha_2, \ldots, \alpha_k$ to be the signed permutation obtained by replacing each $\mu(i)$ with a block that has the pattern α_i if $\mu(i)$ is positive, and α_i^{RC} if $\mu(i)$ is negative. The relative ordering of the blocks is determined by the relative unsigned ordering on μ (meaning we consider simply absolute values). We denote this inflation by $\mu[\alpha_1, \alpha_2, \ldots, \alpha_k]$. For example, $3\bar{1}2[\bar{2}1, \bar{3}1\bar{2}, 1] = \bar{6}5\,2\bar{1}3\,4$.

In the following definition, 1 stands for the identity permutation of length one and $\mathbb{1}_t$ stands for the identity permutation of length t. We write 1^l to denote a sequence of l ones.

Definition 5.3 *We say that $\mu \in B_l$ is valid with respect to $[x_s, x_t]_\mathcal{P}$ if $\mathbb{1}_{k-t+s+1}[1^s, \mu', 1^{k-t}] \in B_k$ is valid with respect to π, where $\mu' = \mu[\mathbb{1}_{m_0}, \mathbb{1}_{m_1}, \ldots, \mathbb{1}_{m_{l-1}}] \in B_{t-s}$ and m_i is the number of irreducible intervals contained in $[a_i, a_{i+1}]$.*

Cohesive partitions play an important role in our proof. In some cases, $\mathcal{P}$ contains a *proper cohesive interval*, which is a cohesive interval $[a_i, a_j]$ where $j \neq i+1$ and $(i, j) \neq (0, l)$. In these cases, we are able to create two new cohesive partitions, $\mathcal{P}' = \{a_0, a_1, \ldots, a_i, a_j, \ldots, a_l\}_<$ and $\widetilde{\mathcal{P}} = \{a_i, a_{i+1}, \ldots, a_j\}_<$, and use them to decompose the action of a valid μ, as described in the following lemma.

Lemma 5.4 *Let $\mu \in B_l$ be valid with respect to $[x_s, x_t]_\mathcal{P}$. If $[a_i, a_j]$ is a proper cohesive interval, there exist $\alpha \in B_{l-j+i+1}$ and $\beta \in B_{j-i}$ such that $\mu = \alpha[1^i, \beta, 1^{l-j}]$. Additionally, β is valid with respect to $[a_i, a_j]_{\widetilde{\mathcal{P}}}$.*

The idea of the proof, which is omitted in this extended abstract, is that there are only two obvious choices for our α, β pair. In one of them, $\alpha(i)$ is positive, and in the other it is negative. One can show that both of these choices make β valid with respect to $[x_i, x_j]_{\widetilde{\mathcal{P}}}$. Since μ is valid with respect to $[x_s, x_t]_\mathcal{P}$, this amounts to examining the behavior of the cycle at the vertices a_i and a_j after applying β; showing that both a_i and a_j have indegree and outdegree one is enough to show that the result is a cycle.

We now have all the tools needed to handle the main theorem. In fact we prove the following more general statement about cohesive partitions.

Proposition 5.5 *Every $\mu \in B_l$ that is valid with respect to $[x_s, x_t]_\mathcal{P}$ transforms $[x_s, x_t]_\mathcal{P}$ by a sequence of valid flips.*

The idea of the proof, again omitted, is that if $\mathcal{P}$ contains a proper cohesive interval $[a_i, a_j]$, we can then use induction to find a sequence of valid flips that first permutes the intervals inside of $[a_i, a_j]$ and

then the intervals outside of $[a_i, a_j]$. The difficulty comes when there is no proper cohesive interval. We deal with this case separately and directly in the proof.

Theorem 4.3, and thus its equivalent restatements Theorems 3.4 and 2.1, follow now as a consequence of Proposition 5.5 taking $l = k$, $s = 0$, $t = k$ and $\mathcal{P} = \mathcal{I}_\pi$, the irreducible partition.

6 Future Work

We have defined a natural equivalence $\sim$ on permutations and characterized the equivalence classes using valid flips. We conjecture that our equivalence classes describe precisely when two permutations are obtained with the same probability in a random walk regardless of the probability distribution on the steps. We are currently working on proving this statement.

Conjecture 6.1 *For $\pi, \tau \in S_n$, $\mathbb{P}(p(X_1, X_2, \ldots, X_{n-1}) = \pi) = \mathbb{P}(p(X_1, X_2, \ldots, X_{n-1}) = \tau)$ for every probability distribution on the i.i.d. random variables $X_1, X_2, \ldots, X_{n-1}$ if and only if $\pi \sim \tau$.*

Acknowledgements

The authors thank Adeline Pierrot for useful ideas during the early stages of this work.

References

[1] J.M. Amigó. *Permutation complexity in dynamical systems*. Springer Series in Synergetics. Springer-Verlag, Berlin, 2010. Ordinal patterns, permutation entropy and all that.

[2] J.M. Amigó, S. Elizalde, and M. Kennel. Forbidden patterns and shift systems. *J. Comb. Theory, Ser. A*, 115(3):485–504, 2008.

[3] K. Archer and S. Elizalde. Cyclic permutations realized by signed shifts. *J. Comb.*, 5(1):1–30, 2014.

[4] C. Bandt, G. Keller, and B. Pompe. Entropy of interval maps via permutations. *Nonlinearity*, 15(5):1595–1602, 2002.

[5] C. Bandt and B. Pompe. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.*, 88:174102, Apr 2002.

[6] C. Bandt and F. Shiha. Order patterns in time series. *J. Time Ser. Anal.*, 28(5):646–665, 2007.

[7] S. Elizalde. The number of permutations realized by a shift. *SIAM J. Discrete Math.*, 23(2):765–786, 2009.

[8] S. Elizalde. Permutations and β-shifts. *J. Combin. Theory Ser. A*, 118(8):2474–2497, 2011.

[9] M. Sinn and K. Keller. Estimation of ordinal pattern probabilities in Gaussian processes with stationary increments. *Comput. Statist. Data Anal.*, 55(4):1781–1790, 2011.

[10] M. Zanin. Forbidden patterns in financial time series. *Chaos*, 18(1):013119, 2008.

[11] L. Zunino, M. Zanin, B. Tabak, D. Pérez, and O. Rosso. Forbidden patterns, permutation entropy and stock market inefficiency. *Physica A*, 388(14):2854–2864, 2009.