# Transcripts and Algebraic Distances in Time Series: Stochastic Properties and Nonparametric Dependence Tests

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

**Christian H. Weiß**

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

**José M. Amigó**

Centro de Investigación Operativa

Universidad Miguel Hernández, Spain

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

# Ordinal Patterns in Time Series

## Introduction

Bandt & Pompe (2002) introduced **ordinal patterns** (OPs) as complexity measures for time series characterized by "*simplicity, extremely fast calculation, robustness, and invariance with respect to nonlinear monotonous transformations*".

**Basic idea** in time-series case: map segments $\boldsymbol{X}_t = (X_t, X_{t+1}, \ldots, X_{t+m-1})$ of length $m$ from continuously distrib., real-valued process $(X_t)_{t\in\mathbb{Z}=\{\ldots,-1,0,1,\ldots\}}$ onto permutations from symmetric group $S_m$ of degree $m$, where selected $\pi_t \in S_m = \{\pi^{[1]}, \ldots, \pi^{[m!]}\}$ expresses order among values in $\boldsymbol{X}_t$ in certain way (see details below).

So original process $(X_t)$ transformed ("discretized") into (symbolic) OP-series $(\pi_t)$ revealing ordinal structure of $(X_t)$:

**marginal distribution** of OP-series $(\pi_t)$ provides insights into **dynamic structure** of original process $(X_t)$.

Expressed as $m!$-dimensional probability vector $\boldsymbol{p}_\pi$ (or frequency vector $\widehat{\boldsymbol{p}}_\pi$ in case of time series data $(x_t)$), with $k$th component being $p_{\pi,k} = P(\pi_t = \pi^{[k]})$.

Analogously, required for asymptotics of $\widehat{\boldsymbol{p}}_\pi$,

$\mathbf{P}_\pi(h) = \big(p_{\pi;kl}(h)\big)$ with $p_{\pi;kl}(h) = P(\pi_t = \pi^{[k]}, \pi_{t+h} = \pi^{[l]})$.

Different (equivalent) approaches to represent OP by permutation from $S_m$, see Berger et al. (2019).

We focus on **permutation representation**.

Then, $\pi = (i_1, \ldots, i_m) \in S_m$ causes ascending order of components of $\boldsymbol{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$, i. e.,

$$x_{i_1} \leq x_{i_2} \leq \ldots \leq x_{i_m}, \quad \text{and} \quad i_{k-1} < i_k \text{ if } x_{i_{k-1}} = x_{i_k}$$

for $k \geq 2$, where "$x_k = x_l$" if ties within $\boldsymbol{x}$. **Here:** Probability of ties $= 0$, so ties at most rarely in data.

**Example:** $(1.2, -0.7, 3.4, 1.9) \quad \mapsto \quad (2, 1, 4, 3)$,

$(1.2, -0.7, 3.4, -0.7) \quad \mapsto \quad (2, 4, 1, 3)$.

Length $m$ of OPs ("$m$-OPs" for short) chosen by user.

If $m = 2$, then downward OP $(2, 1)$ and upward OP $(1, 2)$,

preserves only little information from original process.

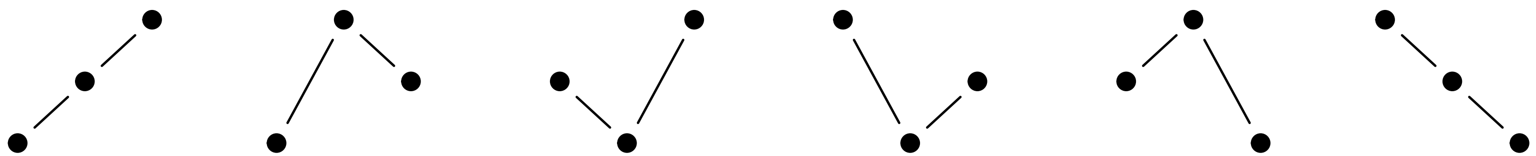
However, range of $\pi_t$ quickly increases with $m$ as $|S_m| = m!$,

so estimation of $\boldsymbol{p}_\pi$ quickly difficult in practice.

Therefore, in time series analysis,

**convenient choice** is $m = 3$ ("3-OPs"; see Bandt, 2019):

$\pi^{[1]} = (1, 2, 3)\quad (1, 3, 2)\quad (2, 1, 3)\quad (2, 3, 1)\quad (3, 1, 2)\quad (3, 2, 1) = \pi^{[6]}$



$\Rightarrow$ sufficiently informative and computationally feasible.

Let $(X_t)$ be *continuously distributed* real-valued process, independent and identically distributed (i. i. d.) *under null*.

Following **properties** crucial for dependence tests:

1. OPs invariant w.r.t. strictly monotonically increasing transformations of $X_t$. Thus, OPs do not depend on actual marginal distribution of $(X_t)$ (**distribution-free** approach).
2. $(X_t)$ is i. i. d. under null ($\rightarrow$ exchangeability). Thus, $\pi_t$ discrete uniform on $S_m$, i. e., $P\big(\pi_t = \pi\big) = 1/m!$ for each $\pi \in S_m$ (**no parameter estimation** required).

**OP-test statistics** built upon $\widehat{\boldsymbol{p}}_\pi$ computed from $\pi_1, \ldots, \pi_n$, where $\pi_t$ from $\boldsymbol{X}_t = (X_t, X_{t+1}, \ldots, X_{t+m-1})$ for $t = 1, 2, \ldots, n$, and where $\boldsymbol{p}_\pi = (1/m!, \ldots, 1/m!)$ under i. i. d.-null:

**Theorem:** (Elsinger, 2010; Weiß, 2022)

$\sqrt{n}\,\big(\widehat{\boldsymbol{p}}_\pi - \boldsymbol{p}_\pi\big) \;\to\; \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_\pi)$ with

$$\boldsymbol{\Sigma}_\pi \;=\; \mathrm{diag}(\boldsymbol{p}_\pi) - \boldsymbol{p}_\pi\,\boldsymbol{p}_\pi^\top \;+\; \sum_{h=1}^{\infty} \Big(\mathbf{P}_\pi(h) + \mathbf{P}_\pi(h)^\top - 2\,\boldsymbol{p}_\pi\,\boldsymbol{p}_\pi^\top\Big).$$

Distribution of OP-test statistics by “Delta method”.

**Nonparametric OP-tests**, suitable for nonlinear dependence, robust against outliers, show appealing power (Weiß, 2022).

HELMUT SCHMIDT UNIVERSITÄT
Universität der Bundeswehr Hamburg

MATH STAT

# Transcripts and Algebraic Distances

Concepts & Properties

Symmetric group $S_m$ forms **algebraic group** together with **composition** "$\circ$" of permutations:

if $\pi_1 = (i_1, i_2, \ldots, i_m)$ and $\pi_2 = (j_1, j_2, \ldots, j_m)$, then

$$\pi_1 \circ \pi_2 = (j_{i_1}, j_{i_2}, \ldots, j_{i_m}).$$

**Transcript** $\tau : S_m \times S_m \to S_m$ maps two OPs $\pi_1, \pi_2 \in S_m$ on OP $\pi \in S_m$ defined by

$$\pi = \tau(\pi_1, \pi_2) := \pi_2 \circ \pi_1^{-1}.$$

Hence, $\pi = \tau(\pi_1, \pi_2)$ transforms $\pi_1$ into $\pi_2$: $\quad \pi \circ \pi_1 = \pi_2$.

By analogy to additive group, where $\tau(x, y) = y - x$,

transcript ≈ "difference"/"dissimilarity" between OPs $\pi_1, \pi_2$.

## Cayley table of transcripts:

| Transcript $\tau$ | $\pi^{[1]}$ | $\pi^{[2]}$ | $\pi^{[3]}$ | $\pi^{[4]}$ | $\pi^{[5]}$ | $\pi^{[6]}$ |
|---|---|---|---|---|---|---|
| $\pi^{[1]} = (1,2,3)$ | $\pi^{[1]}$ | $\pi^{[2]}$ | $\pi^{[3]}$ | $\pi^{[4]}$ | $\pi^{[5]}$ | $\pi^{[6]}$ |
| $\pi^{[2]} = (1,3,2)$ | $\pi^{[2]}$ | $\pi^{[1]}$ | $\pi^{[5]}$ | $\pi^{[6]}$ | $\pi^{[3]}$ | $\pi^{[4]}$ |
| $\pi^{[3]} = (2,1,3)$ | $\pi^{[3]}$ | $\pi^{[4]}$ | $\pi^{[1]}$ | $\pi^{[2]}$ | $\pi^{[6]}$ | $\pi^{[5]}$ |
| $\pi^{[4]} = (2,3,1)$ | $\pi^{[5]}$ | $\pi^{[6]}$ | $\pi^{[2]}$ | $\pi^{[1]}$ | $\pi^{[4]}$ | $\pi^{[3]}$ |
| $\pi^{[5]} = (3,1,2)$ | $\pi^{[4]}$ | $\pi^{[3]}$ | $\pi^{[6]}$ | $\pi^{[5]}$ | $\pi^{[1]}$ | $\pi^{[2]}$ |
| $\pi^{[6]} = (3,2,1)$ | $\pi^{[6]}$ | $\pi^{[5]}$ | $\pi^{[4]}$ | $\pi^{[3]}$ | $\pi^{[2]}$ | $\pi^{[1]}$ |

**Transcript** $\approx$ "difference"/"dissimilarity" between OPs $\pi_1, \pi_2$

$\Rightarrow$ natural relation to **distances** between permutations:

- **Cayley distance** $d_C(\pi_1, \pi_2)$ = minimum number of transpositions to transform $\pi_1$ into $\pi_2$;
- **Kendall distance** $d_K(\pi_1, \pi_2)$ = minimum number of adjacent transpositions to transform $\pi_1$ into $\pi_2$.

$\tau = \pi^{[1]} \;\Rightarrow\; d_C = 0,\ d_K = 0;$ $\quad \tau = \pi^{[2]} \;\Rightarrow\; d_C = 1,\ d_K = 1;$

$\tau = \pi^{[3]} \;\Rightarrow\; d_C = 1,\ d_K = 1;$ $\quad \tau = \pi^{[4]} \;\Rightarrow\; d_C = 2,\ d_K = 2;$

$\tau = \pi^{[5]} \;\Rightarrow\; d_C = 2,\ d_K = 2;$ $\quad \tau = \pi^{[6]} \;\Rightarrow\; d_C = 1,\ d_K = 3;$

see Amigó & Dale (2026) for these and further concepts.

# Transcripts and Algebraic Distances

In what follows, given process $(X_t)$,

we first compute OP-series $(\pi_t)$ with length $m = 3$,

then **transcript series** $(\tau_t)$ with $\tau_t = \pi_{t+1} \circ \pi_t^{-1}$

and corresponding **distance series** $(d_{\mathsf{C},t})$ and $(d_{\mathsf{K},t})$.

Recall that successive 3-OPs "overlap", as

$\pi_t$ refers to $(X_t, X_{t+1}, X_{t+2})$ and $\pi_{t+1}$ to $(X_{t+1}, X_{t+2}, X_{t+3})$.

Also impossible transitions between successive 3-OPs,

see de Sousa & Hlinka (2022).

**Example:** $\pi^{[1]} = (1, 2, 3)$ never followed immediately (lag 1)

by $\pi^{[6]} = (3, 2, 1)$, would only be possible with lag $\geq 2$.

# Transcripts and Algebraic Distances



Altogether, the $m! = 6$ transcripts correspond to following **partition of** $4! = 24$ **4-OPs:**

| $\tau_t = \pi^{[1]}$ | $\tau_t = \pi^{[2]}$ | $\tau_t = \pi^{[3]}$ | $\tau_t = \pi^{[4]}$ | $\tau_t = \pi^{[5]}$ | $\tau_t = \pi^{[6]}$ |
|---|---|---|---|---|---|
| $(1,2,3,4)$ | $(1,2,4,3)$ | $(2,1,3,4)$ | $(1,4,3,2)$ | $(1,3,2,4)$ | $(1,3,4,2)$ |
| $(4,3,2,1)$ | $(4,3,1,2)$ | $(3,4,2,1)$ | $(2,1,4,3)$ | $(1,4,2,3)$ | $(2,4,3,1)$ |
| | | | $(2,4,1,3)$ | $(2,3,1,4)$ | $(3,1,2,4)$ |
| | | | $(3,2,1,4)$ | $(2,3,4,1)$ | $(4,2,1,3)$ |
| | | | $(3,2,4,1)$ | $(3,1,4,2)$ | |
| | | | $(4,1,3,2)$ | $(3,4,1,2)$ | |
| | | | $(4,2,3,1)$ | $(4,1,2,3)$ | |

Therefore, marginal distribution of transcripts computed via marginal distribution of 4-OPs.

**Notations:** $\boldsymbol{p}_\tau = (p_{\tau;1}, \ldots, p_{\tau;m!})^\top$ with $p_{\tau;k} = P(\tau_t = \pi^{[k]})$, $\mathbf{P}_\tau(h) = \big(p_{\tau;kl}(h)\big)$ with $p_{\tau;kl}(h) = P(\tau_t = \pi^{[k]}, \tau_{t+h} = \pi^{[l]})$.

Distribution of distance series via transformation matrices

$$\mathbf{T}_{\mathsf{C}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}, \qquad \mathbf{T}_{\mathsf{K}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Namely,

$$\boldsymbol{p}_{\mathsf{C}} = \mathbf{T}_{\mathsf{C}}\, \boldsymbol{p}_\tau, \qquad \boldsymbol{p}_{\mathsf{K}} = \mathbf{T}_{\mathsf{K}}\, \boldsymbol{p}_\tau,$$

and

$$\mathbf{P}_{\mathsf{C}}(h) = \mathbf{T}_{\mathsf{C}}\, \mathbf{P}_\tau(h)\, \mathbf{T}_{\mathsf{C}}^\top, \qquad \mathbf{P}_{\mathsf{K}}(h) = \mathbf{T}_{\mathsf{K}}\, \mathbf{P}_\tau(h)\, \mathbf{T}_{\mathsf{K}}^\top.$$

**Proposition 3.1.1:** If $(X_t)$ i. i. d.,

then transcripts' marginal distribution

$$\boldsymbol{p}_{\mathcal{T}} = \tfrac{1}{24}\big(2, 2, 2, 7, 7, 4\big)^\top.$$

Cayley and Kendall distances of successive OPs satisfy

$$\boldsymbol{p}_{\mathsf{C}} = \tfrac{1}{12}\big(1, 4, 7\big)^\top \quad \text{with mean } \mu_{\mathsf{C}} = \tfrac{3}{2}, \text{ variance } \sigma_{\mathsf{C}}^2 = \tfrac{5}{12},$$

$$\boldsymbol{p}_{\mathsf{K}} = \tfrac{1}{12}\big(1, 2, 7, 2\big)^\top \quad \text{with mean } \mu_{\mathsf{K}} = \tfrac{11}{6}, \text{ variance } \sigma_{\mathsf{K}}^2 = \tfrac{23}{36}.$$

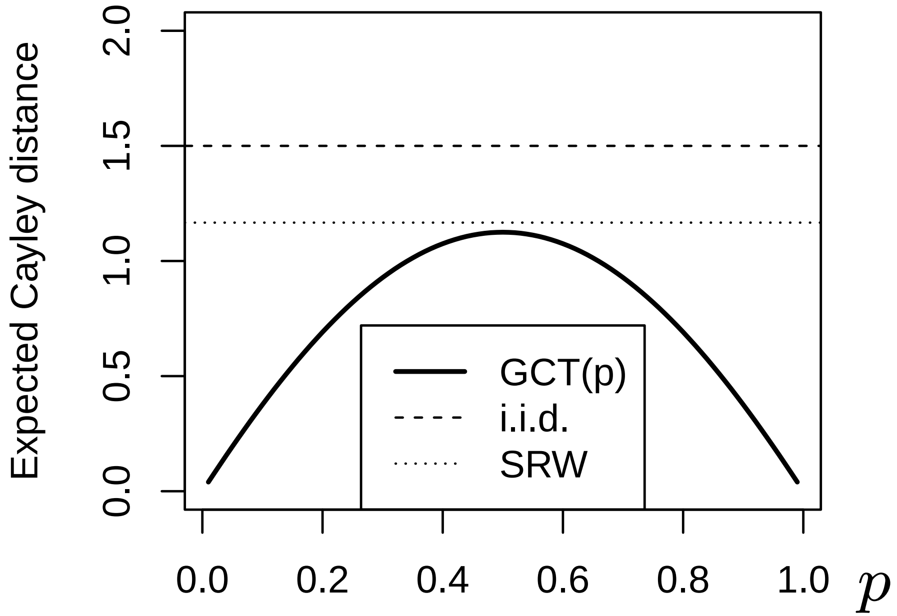
**Closed-form results** also for

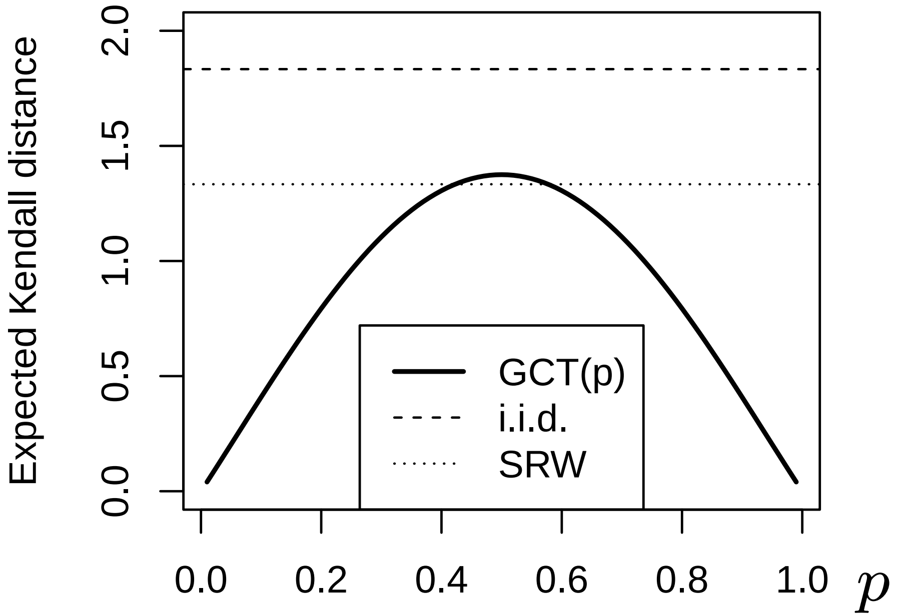
- *symmetric random walk* of de Sousa & Hlinka (2022),
- *generalized coin-tossing* process of Silbernagel & W. (2026).

# Transcripts and Algebraic Distances

**Idea:** use transcripts/distances to test null "$(X_t)$ is i. i. d.".

Seems **promising approach**, for example:



Or, simulating later **alternative scenarios**,

- first-order autoregressive (AR(1)) process $X_t = \phi X_{t-1} + \epsilon_t$ with $\phi \in \{-0.5, 0.5\}$ and i. i. d. normal $\epsilon_t \sim \mathrm{N}(0,1)$;
- first-order quadratic moving-average (QMA(1)) process $X_t = \epsilon_t + 0.8\, \epsilon_{t-1}^2$ with i. i. d. $\epsilon_t \sim \mathrm{N}(0,1)$;
- first-order transposed exponential AR (TEAR(1)) process $X_t = B_t\, X_{t-1} + 0.85\, \epsilon_t$ with i. i. d. exponential $\epsilon_t \sim \mathrm{Exp}(1)$ and Bernoulli $B_t$ with $P(B_t = 1) = 0.15$;
- first-order AR conditional heteroscedasticity ARCH(1) process $X_t = \epsilon_t \sqrt{0.2 + 0.8\, X_{t-1}^2}$ with i. i. d. $\epsilon_t \sim \mathrm{N}(0,1)$; (...)

(. . . ) then again clear deviations from i. i. d.-case:

| Process | $p_{\mathcal{T};1}$ | $p_{\mathcal{T};2}$ | $p_{\mathcal{T};3}$ | $p_{\mathcal{T};4}$ | $p_{\mathcal{T};5}$ | $p_{\mathcal{T};6}$ |
|---|---|---|---|---|---|---|
| *i. i. d.* | *0.083* | *0.083* | *0.083* | *0.292* | *0.292* | *0.167* |
| AR(1), $+0.5$ | 0.150 | 0.115 | 0.115 | 0.256 | 0.254 | 0.110 |
| AR(1), $-0.5$ | 0.041 | 0.042 | 0.042 | 0.293 | 0.293 | 0.287 |
| QMA(1) | 0.103 | 0.076 | 0.094 | 0.348 | 0.220 | 0.160 |
| TEAR(1) | 0.101 | 0.085 | 0.089 | 0.233 | 0.341 | 0.152 |
| ARCH(1) | 0.100 | 0.085 | 0.085 | 0.279 | 0.279 | 0.172 |

Similar deviations for Cayley and Kendall distances.

For asymptotic implementation of dependence tests,

bivariate lag-$h$ distributions under i. i. d.-null required: (. . . )

Transcripts $(\tau_t, \tau_{t+1})$ with lag 1 corresponds to occurrence of certain 5-OPs, which allow to conclude:

**3.2.1 Proposition:** If $(X_t)$ i. i. d., then

$$\mathbf{P}_\tau(1) = \frac{1}{120}\begin{pmatrix} 2 & 1 & 1 & 3 & 3 & 0 \\ 1 & 0 & 1 & 6 & 1 & 1 \\ 1 & 1 & 0 & 1 & 6 & 1 \\ 3 & 1 & 6 & 14 & 4 & 7 \\ 3 & 6 & 1 & 4 & 14 & 7 \\ 0 & 1 & 1 & 7 & 7 & 4 \end{pmatrix},$$

$$\mathbf{P}_\mathrm{C}(1) = \frac{1}{60}\begin{pmatrix} 1 & 1 & 3 \\ 1 & 5 & 14 \\ 3 & 14 & 18 \end{pmatrix} \quad \text{with acf } \rho_\mathrm{C}(1) = -\tfrac{2}{25},$$

$$\mathbf{P}_\mathrm{K}(1) = \frac{1}{60}\begin{pmatrix} 1 & 1 & 3 & 0 \\ 1 & 1 & 7 & 1 \\ 3 & 7 & 18 & 7 \\ 0 & 1 & 7 & 2 \end{pmatrix} \quad \text{with acf } \rho_\mathrm{K}(1) = \tfrac{22}{115}.$$

Transcripts $(\tau_t, \tau_{t+2})$ with lag 2 corresponds to occurrence of certain 6-OPs, which allow to conclude:

**3.2.3 Proposition:** If $(X_t)$ i. i. d., then

$$\mathbf{P}_\tau(2) \;=\; \frac{1}{720}\begin{pmatrix} 2 & 7 & 7 & 18 & 18 & 8 \\ 7 & 0 & 14 & 25 & 6 & 8 \\ 7 & 14 & 0 & 6 & 25 & 8 \\ 18 & 6 & 25 & 77 & 48 & 36 \\ 18 & 25 & 6 & 48 & 77 & 36 \\ 8 & 8 & 8 & 36 & 36 & 24 \end{pmatrix},$$

$$\mathbf{P}_{\mathsf{C}}(2) \;=\; \frac{1}{360}\begin{pmatrix} 1 & 11 & 18 \\ 11 & 42 & 67 \\ 18 & 67 & 125 \end{pmatrix} \qquad \text{with acf } \rho_{\mathsf{C}}(2) = 0,$$

$$\mathbf{P}_{\mathsf{K}}(2) \;=\; \frac{1}{360}\begin{pmatrix} 1 & 7 & 18 & 4 \\ 7 & 14 & 31 & 8 \\ 18 & 31 & 125 & 36 \\ 4 & 8 & 36 & 12 \end{pmatrix} \qquad \text{with acf } \rho_{\mathsf{K}}(2) = \tfrac{8}{115}.$$

Transcripts $(\tau_t, \tau_{t+3})$ with lag 3 corresponds to occurrence of certain 7-OPs, which allow to conclude:

**3.2.5 Proposition:** If $(X_t)$ i. i. d., then

$$\mathbf{P}_\tau(3) = \frac{1}{5040}\begin{pmatrix} 42 & 35 & 35 & 119 & 119 & 70 \\ 35 & 40 & 30 & 116 & 129 & 70 \\ 35 & 30 & 40 & 129 & 116 & 70 \\ 119 & 129 & 116 & 421 & 440 & 245 \\ 119 & 116 & 129 & 440 & 421 & 245 \\ 70 & 70 & 70 & 245 & 245 & 140 \end{pmatrix},$$

$$\mathbf{P}_{\mathsf{C}}(3) = \frac{1}{360}\begin{pmatrix} 3 & 10 & 17 \\ 10 & 40 & 70 \\ 17 & 70 & 123 \end{pmatrix} \text{ with acf } \rho_{\mathsf{C}}(3) = \tfrac{1}{75},$$

$$\mathbf{P}_{\mathsf{K}}(3) = \frac{1}{360}\begin{pmatrix} 3 & 5 & 17 & 5 \\ 5 & 10 & 35 & 10 \\ 17 & 35 & 123 & 35 \\ 5 & 10 & 35 & 10 \end{pmatrix} \text{ with acf } \rho_{\mathsf{K}}(3) = \tfrac{1}{115}.$$

# Transcripts and Algebraic Distances



If $m$-OPs $(\pi_t)$ arise from i. i. d. process $(X_t)$,

then $m$-OPs $(\pi_t)$ known to be $(m-1)$-dependent.

Transcripts derived from 3-OPs correspond to (sets of) 4-OPs

$\Rightarrow$ transcript series $(\tau_t)$ is **3-dependent**.

(same holds for distance series $(d_{\mathsf{C},t})$ and $(d_{\mathsf{K},t})$)

Therefore, sufficient to derive $\mathbf{P}_\tau(1)$, $\mathbf{P}_\tau(2)$, and $\mathbf{P}_\tau(3)$.

Due to 3-dependence of $(\tau_t)$, it holds that

$\mathbf{P}_\tau(h) = \boldsymbol{p}_\tau\,\boldsymbol{p}_\tau^\top$ for lag $|h| \geq 4$ if $(X_t)$ is i. i. d.

Helmut Schmidt Universität
Universität der Bundeswehr Hamburg

MATH STAT

# Nonparametric Tests based on Transcripts and Algebraic Distances

## Asymptotics & Performance

**4.1 Theorem:** Let transcript series $(\tau_t)$ be stationary and $\alpha$-mixing with mixing coefficients $\alpha_i \geq 0$, $i \in \mathbb{N}_0$, satisfying $\sum_{i=0}^{\infty} \alpha_i < \infty$. Then, as $n \to \infty$,

$$\sqrt{n}\,\big(\hat{\boldsymbol{p}}_{\mathcal{T}} - \boldsymbol{p}_{\mathcal{T}}\big) \ \stackrel{d}{\to}\ \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathcal{T}}),$$

where asymptotic covariance matrix $\boldsymbol{\Sigma}_{\mathcal{T}}$ equals

$$\boldsymbol{\Sigma}_{\mathcal{T}} \ = \ \mathrm{diag}(\boldsymbol{p}_{\mathcal{T}}) - \boldsymbol{p}_{\mathcal{T}}\,\boldsymbol{p}_{\mathcal{T}}^{\top} \ + \ \sum_{h=1}^{\infty} \Big(\mathbf{P}_{\mathcal{T}}(h) + \mathbf{P}_{\mathcal{T}}(h)^{\top} - 2\,\boldsymbol{p}_{\mathcal{T}}\,\boldsymbol{p}_{\mathcal{T}}^{\top}\Big).$$

Recall that if $(X_t)$ is i. i. d.,

then $(\tau_t)$ is “only” 3-dependent. Thus, $(\ldots)$

**4.2 Corollary:** If $(X_t)$ i. i. d., then $\boldsymbol{p}_\mathcal{T} = \frac{1}{24}\big(2,2,2,7,7,4\big)^\top$, and as $n\to\infty$,

$$\sqrt{n}\,\big(\hat{\boldsymbol{p}}_\mathcal{T} - \boldsymbol{p}_\mathcal{T}\big) \ \xrightarrow{d}\ \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_\mathcal{T}),$$

where asymptotic covariance matrix $\boldsymbol{\Sigma}_\mathcal{T}$ has rank 5:

$$\boldsymbol{\Sigma}_\mathcal{T} \ = \ \frac{1}{20\,160}\begin{pmatrix} 1820 & 28 & 28 & -462 & -462 & -952 \\ 28 & 1020 & 380 & -406 & -406 & -616 \\ 28 & 380 & 1020 & -406 & -406 & -616 \\ -462 & -406 & -406 & 6259 & -4453 & -532 \\ -462 & -406 & -406 & -4453 & 6259 & -532 \\ -952 & -616 & -616 & -532 & -532 & 3248 \end{pmatrix}.$$

**4.3 Corollary:** If $(X_t)$ i. i. d., then $\boldsymbol{p}_{\mathsf{C}} = \frac{1}{12}\big(1,4,7\big)^\top$ and $\boldsymbol{p}_{\mathsf{K}} = \frac{1}{12}\big(1,2,7,2\big)^\top$, and as $n \to \infty$,

$$\sqrt{n}\,\big(\hat{\boldsymbol{p}}_{\mathsf{C}} - \boldsymbol{p}_{\mathsf{C}}\big) \stackrel{d}{\to} \mathsf{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathsf{C}}) \quad \text{and} \quad \sqrt{n}\,\big(\hat{\boldsymbol{p}}_{\mathsf{K}} - \boldsymbol{p}_{\mathsf{K}}\big) \stackrel{d}{\to} \mathsf{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\mathsf{K}}),$$

where

$$\boldsymbol{\Sigma}_{\mathsf{C}} = \frac{1}{720}\begin{pmatrix} 65 & -32 & -33 \\ -32 & 128 & -96 \\ -33 & -96 & 129 \end{pmatrix} \quad \text{with rank 2,}$$

$$\boldsymbol{\Sigma}_{\mathsf{K}} = \frac{1}{720}\begin{pmatrix} 65 & 2 & -33 & -34 \\ 2 & 100 & -58 & -44 \\ -33 & -58 & 129 & -38 \\ -34 & -44 & -38 & 116 \end{pmatrix} \quad \text{with rank 3.}$$

**4.4 Corollary:** If $(X_t)$ i. i. d., then $\mu_{\mathsf{C}} = \frac{3}{2}$ and $\mu_{\mathsf{K}} = \frac{11}{6}$, and as $n \to \infty$, mean distances $\overline{d_{\mathsf{C}}}$ and $\overline{d_{\mathsf{K}}}$ satisfy

$$\sqrt{n}\left(\overline{d_{\mathsf{C}}} - \mu_{\mathsf{C}}\right) \overset{d}{\to} \mathrm{N}\big(0, \tfrac{13}{36}\big) \quad \text{and} \quad \sqrt{n}\left(\overline{d_{\mathsf{K}}} - \mu_{\mathsf{K}}\right) \overset{d}{\to} \mathrm{N}\big(0, \tfrac{59}{60}\big).$$

We obtain various **nonparametric dependence tests:**

$\overline{d_{\mathsf{C}}}$- **and** $\overline{d_{\mathsf{K}}}$**-test** with (two-sided) critical or P-values,

and upper-sided **entropy-like tests** based on

$$H(\hat{\boldsymbol{p}}) = -\sum_{i=1}^{k} \hat{p}_i \left(\ln \hat{p}_i - \ln p_i^{(0)}\right),$$

$$\Delta(\hat{\boldsymbol{p}}) = \left(\hat{\boldsymbol{p}} - \boldsymbol{p}^{(0)}\right)^{\top} \operatorname{diag}\left(\boldsymbol{p}^{(0)}\right)^{-1} \left(\hat{\boldsymbol{p}} - \boldsymbol{p}^{(0)}\right).$$

**4.5 Proposition:** If $(X_t)$ i. i. d., then

(a) $-2n\,H_\tau\big(\hat{\boldsymbol{p}}_\tau\big)$ and $n\,\Delta_\tau\big(\hat{\boldsymbol{p}}_\tau\big)$ asymptotically $\sim\ \sum_{i=1}^{5} \lambda_{\tau;i}\,\chi^2_{1;i}$ with

$\lambda_{\tau;1} = \frac{1339}{735},\ \lambda_{\tau;2} \approx 1.547,\ \lambda_{\tau;3} \approx 0.926,\ \lambda_{\tau;4} \approx 0.718,\ \lambda_{\tau;5} = \frac{8}{21};$

(b) $-2n\,H_{\mathsf{C}}\big(\hat{\boldsymbol{p}}_{\mathsf{C}}\big)$ and $n\,\Delta_{\mathsf{C}}\big(\hat{\boldsymbol{p}}_{\mathsf{C}}\big)$ asymptotically $\sim\ \sum_{i=1}^{2} \lambda_{\mathsf{C};i}\,\chi^2_{1;i}$ with

$$\lambda_{\mathsf{C};1} = \frac{6}{5}, \qquad \lambda_{\mathsf{C};2} = \frac{76}{105};$$

(c) $-2n\,H_{\mathsf{K}}\big(\hat{\boldsymbol{p}}_{\mathsf{K}}\big)$ and $n\,\Delta_{\mathsf{K}}\big(\hat{\boldsymbol{p}}_{\mathsf{K}}\big)$ asymptotically $\sim\ \sum_{i=1}^{3} \lambda_{\mathsf{K};i}\,\chi^2_{1;i}$ with

$$\lambda_{\mathsf{K};1} \approx 1.5468, \qquad \lambda_{\mathsf{K};2} \approx 0.9260, \qquad \lambda_{\mathsf{K};3} \approx 0.7177.$$

**In our paper,**

- comprehensive power simulations with above alternatives;
- comparison to former OP-tests by Weiß (2022);
- illustrative data example on monthly mean air temperature (in °C) at "Hamburg–Fuhlsbüttel" (Germany) for 1936–2024.

**Quick summary:**

$H_\mathcal{T}$- and $\Delta_\mathcal{T}$-tests with appealing power for most alternatives. For AR(1), even surpassed by $\overline{d}_\mathsf{K}$-test.
$\overline{d}_\mathsf{C}$-test only excelled for ARCH(1). With exception of TEAR(1), novel transcript-tests *more powerful than former OP-tests*, and often also outperform acf.

# Conclusions & Outlook



Transcripts and algebraic distances based on OPs,
stochastic properties and asymptotics of test statistics,
nonparametric dependence tests with appealing power.

**Work in progress & future research:**

- control charts based on transcripts and algebraic distances, in analogy to Weiß & Testik (2023);
- adapt to generalized OPs of Weiß & Schnurr (2024) for discrete-valued processes, or to
- spatial OPs like in Weiß & Kim (2024), or to
- cross-dependence in multiv. t. s. (Silbernagel et al., 2025).

# Thank You

# for Your Interest!

**Christian H. Weiß**

Department of Mathematics & Statistics

Helmut Schmidt University, Hamburg

`weissc@hsu-hh.de`

# Literature


Amigó & Dale (2025) Permutation-based distances . . . *Entropy* **27**, 913.

Bandt (2019) Small order patterns . . . *Entropy* **21**, 613.

Bandt & Pompe (2002) Permutation entropy . . . *Phys Rev L* **88**, 174102.

Berger et al. (2019) Teaching ordinal patterns . . . *Entropy* **21**, 1023.

Elsinger (2010) Independence tests . . . *Working paper* **165**, Öst. Nat.bank.

Silbernagel et al. (2025) Non-par. tests for cross-dep. . . *CSDA* **210**, 108189.

Silbernagel & Weiß (2026) The joint asymptotic . . . *Chaos* **36**, 023145.

de Sousa & Hlinka (2022) Assessing serial depend. . . *Chaos* **32**, 073126.

Weiß (2022) Non-parametric tests . . . *Chaos* **32**, 093107.

Weiß & Kim (2024) Using spatial ordinal . . . *Spat Stat* **59**, 100800.

Weiß & Schnurr (2024) Generalized ordinal . . . *J Nonpar S* **36**, 573–599.

Weiß & Testik (2023) Non-param. control ch. . . *Technomet* **65**, 340–350.