

Discrete entropies of orthogonal polynomials

A.I. Aptekarev J.S. Dehesa A. Martínez-Finkelshtein* R. Yáñez

February 2, 2008

Abstract

Let p_n , $n \in \mathbb{N}$, be the n th orthonormal polynomial on \mathbb{R} , whose zeros are $\lambda_j^{(n)}$, $j = 1, \dots, n$. Then for each $j = 1, \dots, n$,

$$\vec{\Psi}_j^2 \stackrel{\text{def}}{=} (\Psi_{1j}^2, \dots, \Psi_{nj}^2)$$

with

$$\Psi_{ij}^2 = p_{i-1}^2(\lambda_j^{(n)}) \left(\sum_{k=0}^{n-1} p_k^2(\lambda_j^{(n)}) \right)^{-1}, \quad i = 1, \dots, n,$$

defines a discrete probability distribution. The Shannon entropy of the sequence $\{p_n\}$ is consequently defined as

$$\mathcal{S}_{n,j} \stackrel{\text{def}}{=} - \sum_{i=1}^n \Psi_{ij}^2 \log(\Psi_{ij}^2).$$

In the case of Chebyshev polynomials of the first and second kinds an explicit and closed formula for $\mathcal{S}_{n,j}$ is obtained, revealing interesting connections with the number theory. Besides, several results of numerical computations exemplifying the behavior of $\mathcal{S}_{n,j}$ for other families are also presented.

AMS MOS Classification: 33C45, 41A58, 42C05, 94A17

Keywords: orthogonal polynomials, Shannon entropy, Chebyshev polynomials, Euler-Maclaurin formula

1 Introduction

Given a probability Borel measure μ supported on the real line \mathbb{R} with infinite number of points of increase, we can build a sequence of orthonormal polynomials $p_n(\lambda) = \kappa_n \lambda^n +$ lower degree terms, $n = 0, 1, 2, \dots$, uniquely determined if all $\kappa_n > 0$, such that

$$\int p_n(\lambda) p_m(\lambda) d\mu(\lambda) = \delta_{nm}, \quad m, n = 0, 1, 2, \dots$$

*Corresponding author.

Beside their importance in approximation theory and multiple branches of applied and pure mathematics, orthogonal polynomials constitute a noteworthy object from the point of view of the information theory. This interest originated in the framework of the modern density functional theory [14, 15, 19], that states that the physical and chemical properties of fermionic systems (atoms, molecules, nuclei, solids) may be completely described by means of the single-particle probability density. For instance, if the solution of the time-independent Schrödinger equation in a D -dimensional position space for an single particle system,

$$H\Psi(\vec{r}) = E\Psi(\vec{r}), \quad \vec{r} = (x_1, \dots, x_D),$$

is the wave function $\Psi(\vec{r})$, then the position density of the system is $\rho(\vec{r}) = |\Psi(\vec{r})|^2$. Analogously, the wave function in momentum space $\widehat{\Psi}(\vec{p})$, which is the Fourier transform of $\Psi(\vec{r})$, gives the momentum density $\gamma(\vec{p}) = |\widehat{\Psi}(\vec{p})|^2$.

Information measures of these densities are closely related to fundamental and experimentally measurable physical quantities, which makes them useful in the study of the structure and dynamics of atomic and molecular systems. For instance, the *Boltzmann-Gibbs-Shannon* (position-space) *entropy*

$$\mathcal{B}(\rho) = - \int \rho(\vec{r}) \log \rho(\vec{r}) d\vec{r} \quad (1)$$

measures the uncertainty in the localization of the particle in space. Lower entropy corresponds to a more concentrated wave function, with smaller uncertainty, and hence, higher accuracy in predicting the localization of the particle. The well known inequality [5, 8, 9]

$$\mathcal{B}(\rho) + \mathcal{B}(\gamma) \geq D(1 + \log \pi) \quad (2)$$

is an expression of the position-momentum uncertainty principle, much stronger than the renowned Heisenberg relation, that plays a major role in quantum mechanics (see [22]).

The study of the information measures of orthogonal polynomials is motivated by the fact that the densities of many quantum mechanical systems with shape-invariant potentials (e.g., the harmonic oscillator and the hydrogenic systems) typically contain terms of the form $p_n^2 \mu'$. Explicit formulas, numerical algorithms and asymptotic behavior have been studied both for the Boltzmann-Gibbs-Shannon (or differential) entropy

$$\mathcal{B}_n = - \int p_n^2(\lambda) \log (p_n^2(\lambda)\mu'(\lambda)) d\mu(\lambda)$$

and for the the relative entropy (or the Kullback-Leibler information)

$$\mathcal{K}_n = - \int p_n^2(\lambda) \log (p_n^2(\lambda)) d\mu(\lambda);$$

see e.g.[3, 4, 7, 10, 11, 12, 13, 20] and the references therein. In particular, it has been shown that for Chebyshev orthonormal polynomials of the first kind the relative entropy \mathcal{K}_n does not depend on n ,

$$\mathcal{K}_n = \log(2) - 1,$$

and that this value is asymptotically maximal among all orthogonality measures on $[-1, 1]$ (see [7]), giving a formal explanation to the intuitive notion that these polynomials are the most “uniformly” distributed ones.

However, there are several discrete measures naturally associated with a sequence of orthogonal polynomials. The analysis of such measures requires the use of the “genuine” entropy studied by Shannon. In order to stress the discrete character of this entropy, hereafter we refer to it as *Shannon entropy* and denote it by the letter \mathcal{S} . The evaluation of Shannon entropy for discrete distributions is a basic question of information theory (see e.g. [17, 18]); unlike for \mathcal{B}_n and \mathcal{K}_n , there are no known results for the Shannon entropy of the orthogonal polynomials related distributions, due in part to the technical difficulties of the explicit evaluation of sums.

It is well known that the orthonormal polynomials p_n satisfy a three-term recurrence relation of the form

$$\lambda p_i(\lambda) = b_{i+1} p_{i+1}(\lambda) + a_{i+1} p_i(\lambda) + b_i p_{i-1}(\lambda), \quad i = 0, 1, \dots, n-2, \quad p_{-1} = 0, p_0 = 1. \quad (3)$$

Using its coefficients we can define the $n \times n$ Jacobi matrix ($n \in \mathbb{N}$),

$$L_n = \begin{pmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_1 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & a_{n-1} & b_{n-1} \\ & & & b_{n-1} & a_n \end{pmatrix}, \quad (4)$$

which determines a self-adjoint linear operator (discrete Schrödinger operator) $L_n : \mathbb{R}^n \rightarrow \mathbb{R}^n$ by

$$L \vec{e}_i = b_i \vec{e}_{i+1} + a_i \vec{e}_i + b_{i-1} \vec{e}_{i-1}, \quad i = 1, \dots, n,$$

where $\vec{e}_1, \dots, \vec{e}_n$ is the canonical basis in \mathbb{R}^n , and we agree that $\vec{e}_0 = \vec{e}_{n+1} = \vec{0}$. Moreover, up to a constant factor, $p_n(\lambda) = \det(L_n - \lambda I)$, which shows that the eigenvalues $\lambda_k^{(n)}$, $k = 1, \dots, n$, are the zeros of p_n , and

$$\vec{P}_k = \left(p_0, p_1(\lambda_k^{(n)}), \dots, p_{n-1}(\lambda_k^{(n)}) \right)^T,$$

are eigenvectors corresponding to different eigenvalues.

Let $\langle \cdot, \cdot \rangle$ denote the standard (euclidean) inner product in \mathbb{R}^n , and

$$\ell_n(\lambda) \stackrel{\text{def}}{=} \left(\sum_{k=0}^{n-1} p_k^2(\lambda) \right)^{-1} = \frac{1}{\langle \vec{P}_k, \vec{P}_k \rangle} \quad (5)$$

the n -th Christoffel function.

If we normalize

$$\vec{\Psi}_k = \sqrt{\ell_n(\lambda_k^{(n)})} \vec{P}_k,$$

then a consequence of the well known Christoffel-Darboux formula is that

$$\langle \vec{\Psi}_i, \vec{\Psi}_j \rangle = \delta_{ij}, \quad i, j = 1, \dots, n. \quad (6)$$

In other words, the $n \times n$ matrix

$$\Psi = \left(\sqrt{\ell_n(\lambda_j^{(n)})} p_{i-1}(\lambda_j^{(n)}) \right)_{i,j=1}^n,$$

made of columns $\vec{\Psi}_j$, $j = 1, \dots, n$, is orthogonal, so that the squares of the components of each (column) vector $\vec{\Psi}_j$, $j = 1, \dots, n$, give a discrete probability distribution, and these distributions are mutually orthogonal in the sense of (6).

Recall that given a probability measure $\mu = (\mu_1, \mu_2, \dots, \mu_n)$ on a system of n points, e. g. $\sum_{j=1}^n \mu_j = 1$, the standard *Shannon entropy* reads $\mathcal{S}(\mu) = -\sum_{j=1}^n \mu_j \log \mu_j$. By Jensen's inequality,

$$0 \leq \mathcal{S}(\mu) \leq \log(n), \quad (7)$$

and the maximum corresponds to a uniform probability distribution. In this sense, it is quite natural to think of the Shannon entropy as a measure of uncertainty.

Remark 1 We can give the following geometric interpretation to the Shannon entropy. Given in \mathbb{R}^n an orthonormal basis $\{\vec{e}_i\}$, any vector $\vec{v} \in \mathbb{R}^n$ has a unique representation

$$\vec{v} = \sum_{i=1}^n \langle \vec{v}, \vec{e}_i \rangle \vec{e}_i.$$

Assume that $\vec{v} \in S^{n-1}$, that is, $\|\vec{v}\| = 1$, where $\|\cdot\|$ means the Euclidean norm. A natural way of measuring a relative distance of \vec{v} from the basis $\{\vec{e}_i\}$ is by means of the Shannon entropy

$$\mathcal{S}_n \stackrel{\text{def}}{=} -\sum_{i=1}^n p_i \log(p_i), \quad p_i \stackrel{\text{def}}{=} \langle \vec{e}_i, \vec{v} \rangle^2, \quad i = 1, \dots, n. \quad (8)$$

Indeed, if $\vec{v} = \vec{e}_k$ for a certain k , then $p_j = \delta_{jk}$, and $\mathcal{S}_n = 0$. On the contrary, if \vec{v} is "equidistant" from all vectors \vec{e}_j 's, then all $p_j = 1/n$, and \mathcal{S}_n attains its maximum, $\mathcal{S}_n = \log(n)$.

Motivated by the discussion above, we introduce the *discrete entropy of orthonormal polynomials* p_n , defined as the Shannon entropy of the probability distribution given by each column of Ψ :

$$\begin{aligned} \mathcal{S}_{n,j} &\stackrel{\text{def}}{=} -\sum_{i=1}^n \Psi_{ij}^2 \log(\Psi_{ij}^2) \\ &= -\ell_n(\lambda_j^{(n)}) \sum_{i=1}^n p_{i-1}^2(\lambda_j^{(n)}) \log\left(\ell_n(\lambda_j^{(n)}) p_{i-1}^2(\lambda_j^{(n)})\right) \\ &= -\log\left(\ell_n(\lambda_j^{(n)})\right) - \ell_n(\lambda_j^{(n)}) \sum_{i=1}^n p_{i-1}^2(\lambda_j^{(n)}) \log\left(p_{i-1}^2(\lambda_j^{(n)})\right), \quad j = 1, \dots, n, \end{aligned} \quad (9)$$

which can be generalized as

$$\mathcal{S}_n(\lambda) \stackrel{\text{def}}{=} -\log(\ell_n(\lambda)) - \ell_n(\lambda) \sum_{i=1}^n p_{i-1}^2(\lambda) \log(p_{i-1}^2(\lambda)), \quad (10)$$

so that $\mathcal{S}_{n,j} = \mathcal{S}_n(\lambda_j^{(n)})$.

Unlike the Boltzmann-Gibbs-Shannon entropy, \mathcal{S}_n does not depend on the weight function μ' , and is suitable both for discrete and continuous orthogonality (cf. some numerical experiments in Section 5).

Remark 2 Since Ψ is an orthogonal matrix, its rows

$$\left(\sqrt{\ell_n(\lambda_1^{(n)})} p_{i-1}(\lambda_1^{(n)}), \dots, \sqrt{\ell_n(\lambda_n^{(n)})} p_{i-1}(\lambda_n^{(n)}) \right)$$

are also orthogonal vectors of \mathbb{R}^n :

$$\delta_{ij} = \sum_{k=1}^n \ell_n(\lambda_k^{(n)}) p_{i-1}(\lambda_k^{(n)}) p_{j-1}(\lambda_k^{(n)}) = \int p_{i-1}(\lambda) p_{j-1}(\lambda) d\mu_n(\lambda),$$

where μ_n is the normalized counting measure of zeros of p_n :

$$\mu_n = \sum_{k=1}^n \ell_n(\lambda_k^{(n)}) \delta_{\lambda_k^{(n)}}.$$

Hence, we may define the dual discrete entropy, corresponding to rows of Ψ :

$$\mathcal{S}_n^i \stackrel{\text{def}}{=} - \sum_{j=1}^n \ell_n(\lambda_j^{(n)}) p_{i-1}^2(\lambda_j^{(n)}) \log \left(\ell_n(\lambda_j^{(n)}) p_{i-1}^2(\lambda_j^{(n)}) \right), \quad j = 1, \dots, n.$$

A basic question of information theory is the evaluation of the Shannon entropy. In this paper we compute explicitly the discrete entropy $\mathcal{S}_{n,j}$ corresponding to Chebyshev orthonormal polynomials of the first and second kinds. A straightforward interpretation of (9) as Riemann sums allows to find the first two terms of the asymptotic expansion of $\mathcal{S}_{n,j}$ for fixed j and large n ; these terms do not depend on j . However, numerical experiments reveal the existence of certain picks, pointing downwards, whose position was not clear a priori (see Figure 2). The formulas presented below give a complete explanation of this phenomenon and exhibit nice connections with relevant objects from the number theory.

In order to state our results we need to introduce an auxiliary function

$$\mathcal{R}(x) \stackrel{\text{def}}{=} x(\Psi(1-x) + 2\gamma + \Psi(1+x)), \quad x \in [0, 1], \quad (11)$$

where γ is the Euler constant, and $\Psi(x) = \Gamma'(x)/\Gamma(x)$ is the digamma function. Alternatively, \mathcal{R} can be given by its Taylor series expansion, absolutely convergent for $|x| < 1$ (cf. formula (6.3.14) in [1]),

$$\mathcal{R}(x) = -2 \sum_{k=1}^{\infty} \zeta(2k+1) x^{2k+1}, \quad (12)$$

where $\zeta(\cdot)$ is the Riemann zeta function.

Recall that *Chebyshev polynomials of the first kind* are given by the explicit formula

$$p_m(\lambda) = T_m(\lambda) = \begin{cases} 1, & \text{if } m = 0, \\ \sqrt{2} \cos(m\theta), & \text{otherwise,} \end{cases} \quad \lambda = \cos \theta. \quad (13)$$

They are orthonormal with respect to the weight

$$w(\lambda) = \frac{1}{\pi} \frac{1}{\sqrt{1-\lambda^2}} \quad \text{on } [-1, 1].$$

Theorem 1 Let $n \in \mathbb{N}$, $j \in \{1, 2, \dots, n\}$. For orthonormal Chebyshev polynomials of the first kind, the discrete entropy has the following expression:

$$\mathcal{S}_{n,j} = \log n + \log 2 - 1 + \frac{\log 2}{n} + \mathcal{R}\left(\frac{d}{2n}\right), \quad d = \text{GCD}(2j - 1, n).$$

Hereafter GCD stands for the greatest common divisor.

Remark 3 Observe that coefficients in the series expansion (12) are all positive, so that $\mathcal{R}(x) < 0$ and is strictly decreasing for $x \in (0, 1)$ (see Figure 1). Since $1 \leq d \leq n$, we see that

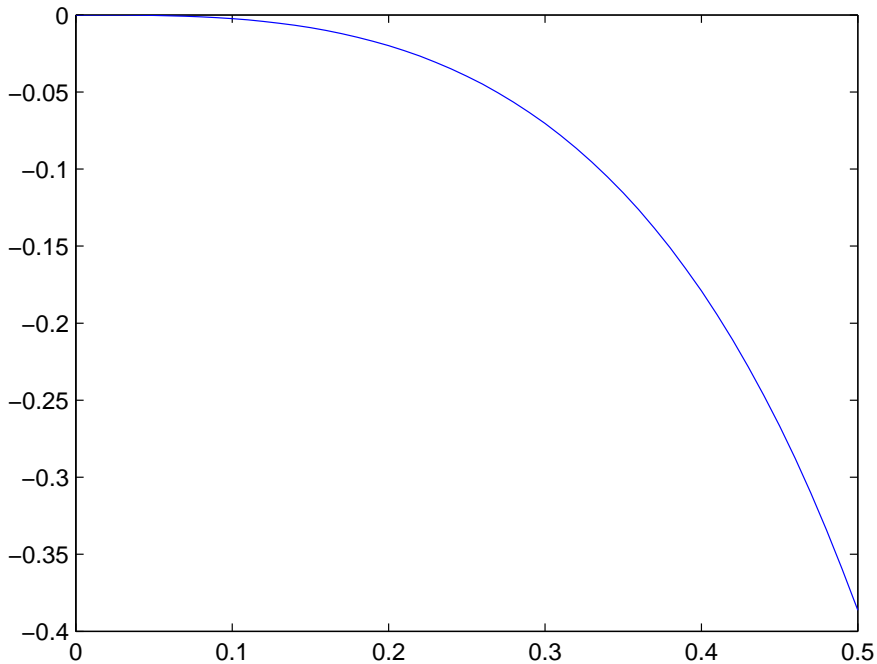


Figure 1: Function $\mathcal{R}(x)$, for $x \in [0, 1/2]$.

$$\max_{j \in \{1, 2, \dots, n\}} \mathcal{S}_{n,j} = \log n + \log 2 - 1 + \frac{\log 2}{n} - 2 \sum_{k=1}^{\infty} \zeta(2k + 1) \left(\frac{1}{2n}\right)^{2k+1},$$

attained when $\text{GCD}(2j - 1, n) = 1$. Furthermore, if n is odd, then $\mathcal{S}_{n,j}$ attains its minimum

$$\min_{j \in \{1, 2, \dots, n\}} \mathcal{S}_{n,j} = \log n - \log 2 + \frac{\log 2}{n},$$

at a single value $j = (n + 1)/2$. It is the only local minimum of $\mathcal{S}_{n,j}$ if $n \geq 3$ is prime.

The reader can compare these observations with the results of numerical experiments shown in Figure 2.

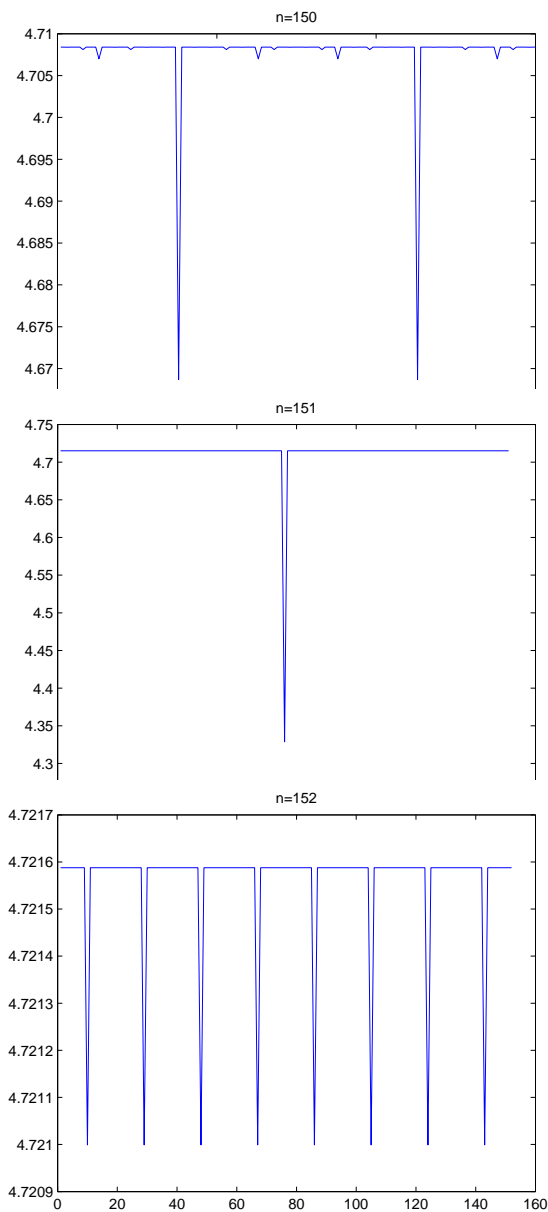


Figure 2: Chebyshev polynomials of the first kind: entropy $S_{n,j}$ for $n = 150, 151$ and 152 .

The *Chebyshev polynomials of the second kind* are

$$p_m(\lambda) = U_m(\lambda) = \frac{\sin[(m+1)\arccos(\lambda)]}{\sqrt{1-\lambda^2}} = \frac{\sin[(m+1)\theta]}{\sin(\theta)}, \quad \lambda = \cos \theta, \quad m \geq 0. \quad (14)$$

They are orthonormal with respect to the weight

$$w(\lambda) = \frac{2}{\pi} \sqrt{1-\lambda^2} \quad \text{on } [-1, 1].$$

Theorem 2 *Let $n \in \mathbb{N}$, $j \in \{1, 2, \dots, n\}$. For orthonormal Chebyshev polynomials of the second kind, the discrete entropy has the following expression:*

$$S_{n,j} = \log(n+1) + \log 2 - 1 + \mathcal{R}\left(\frac{d}{n+1}\right), \quad d = \text{GCD}(j, n+1).$$

Remark 4 Since $1 \leq d \leq n$, we see that

$$\max_{j \in \{1, 2, \dots, n\}} \mathcal{S}_{n,j} = \log(n+1) + \log 2 - 1 + \mathcal{R}\left(\frac{1}{n+1}\right),$$

attained when $\text{GCD}(j, n+1) = 1$. Furthermore, if n is odd, $\mathcal{S}_{n,j}$ attains its minimum

$$\min_{j \in \{1, 2, \dots, n\}} \mathcal{S}_{n,j} = \log\left(\frac{n+1}{2}\right),$$

at a single value $j = (n+1)/2$. It is the only local minimum of $\mathcal{S}_{n,j}$ if $n \geq 3$ is prime.

Remark 5 The leading term $\log n$ in both cases shows that the values $p_0^2, p_1^2(\lambda_k^{(n)}), \dots, p_{n-1}^2(\lambda_k^{(n)})$, normalized by an appropriate factor, are approximately equidistributed. Comparing formulas from Theorem 1 and 2 we see that unlike for the Boltzmann entropy, the discrete entropy of the Chebyshev polynomials of the first kind is generally smaller.

The rest of the article is organized as follows. In the next section (that might have an independent interest) we discuss some piece-wise linear endomorphisms of \mathbb{R} and their connection with permutations. This allows to reduce the analysis of the general discrete entropy $\mathcal{S}_{n,j}$ to some specific values of the index j . A modification of the Euler-Maclaurin summation formula is the key to the proof of Theorem 1 in Section 3. The close connection between polynomials of the first and second kinds allows us to avoid similar cumbersome computations in the proof of Theorem 2 in Section 4. Finally, we discuss some numerical result obtained for the discrete entropy for other important families of orthogonal polynomials.

2 Piece-wise linear endomorphism of \mathbb{R} and permutations

The key role is played by the following auxiliary function:

Definition 3 For each pair of values $n, j \in \mathbb{N}$, let $\varphi_j^{(n)}(x)$ denote the linear spline on \mathbb{R} with nodes at $\{mn/j\}_{m \in \mathbb{Z}}$ interpolating the values

$$\left\{ \left(\frac{mn}{j}, n \frac{1 - (-1)^m}{2} \right) \right\}_{m \in \mathbb{Z}}.$$

Functions $\varphi_j^{(n)}$ can be explicitly described by

$$x \in \left[\frac{2k-1}{j}n, \frac{2k+1}{j}n \right] \text{ for } k \in \mathbb{Z} \quad \Rightarrow \quad \varphi_j^{(n)}(x) = |jx - 2kn| \quad (15)$$

(see Figure 3).

In order to summarize necessary properties of functions $\varphi_j^{(n)}$ we need to introduce some notation. We denote by $a \equiv b \pmod{c}$ the standard arithmetic congruence of a and b modulo c , $\text{GCD}(a, b)$ stands for the greatest common divisor of integer numbers a and b , and $\mathbb{N}_0 \stackrel{\text{def}}{=} \mathbb{N} \cup \{0\}$. We define also the remainder function $\mathcal{D} : \mathbb{N}_0 \times \mathbb{N} \rightarrow \mathbb{Z}$, by

$$\mathcal{D}(p, q) = r \quad \text{if and only if} \quad -q/2 \leq r < q/2 \text{ and } p \equiv r \pmod{q}$$

(note that this definition is shifted with respect to a standard concept of remainder; thus, $\mathcal{D}(p, q)$ takes also negative values).

Main Lemma Let $n, j \in \mathbb{N}$, with $\text{GCD}(j, n) = d$.

(i) We have

$$\varphi_j^{(n)}\left(x + \frac{n}{d}\right) = \begin{cases} n - \varphi_j^{(n)}(x), & \text{if } j/d \text{ is odd,} \\ \varphi_j^{(n)}(x), & \text{if } j/d \text{ is even,} \end{cases} \quad x \in \mathbb{R}, \quad (16)$$

and for $k \in \mathbb{Z}$,

$$\varphi_j^{(n)}(k) = d \cdot \varphi_{j/d}^{(n/d)}(k). \quad (17)$$

(ii) If $d = n$, then $\varphi_j^{(n)}(\mathbb{Z}) \subset \{0, n\}$.

(iii) If $\text{GCD}(j, 2n) = d < n$, then

$$\varphi_j^{(n)}(\{1, \dots, n-1\} \setminus \{0, n\}) = \left\{ dm : m = 1, \dots, \frac{n}{d} - 1 \right\}, \quad (18)$$

and for any $m \in \{1, \dots, \frac{n}{d} - 1\}$,

$$\text{card}\{k \in \{1, \dots, n-1\} : \varphi_j^{(n)}(k) = dm\} = d. \quad (19)$$

(iv) If $\text{GCD}(j, 2n) = 2d$ and $d < n$, then

$$\varphi_j^{(n)}(\{1, \dots, n-1\} \setminus \{0, n\}) = \left\{ 2dm : m = 1, \dots, \frac{n-d}{2d} \right\}, \quad (20)$$

and for any $m \in \{1, \dots, \frac{n-d}{2d}\}$,

$$\text{card}\{k \in \{1, \dots, n-1\} : \varphi_j^{(n)}(k) = 2dm\} = 2d. \quad (21)$$

We prove this lemma establishing a number of intermediate auxiliary results.

Proposition 4 Let $n, j \in \mathbb{N}$. Function $\varphi_j^{(n)} : \mathbb{R} \rightarrow [0, n]$ satisfies:

(i) $\varphi_j^{(n)}$ is even: $\varphi_j^{(n)}(-x) = \varphi_j^{(n)}(x)$, $x \in \mathbb{R}$.

(ii) Symmetry:

$$\varphi_j^{(n)}\left(x + \frac{n}{j}\right) = n - \varphi_j^{(n)}(x), \quad x \in \mathbb{R}. \quad (22)$$

In particular, $\varphi_j^{(n)}$ is periodic with period $2n/j$.

(iii) For every $x \in \mathbb{R}$, either

$$\frac{\varphi_j^{(n)}(x) - jx}{2n} \in \mathbb{Z} \quad \text{or} \quad \frac{\varphi_j^{(n)}(x) + jx}{2n} \in \mathbb{Z}. \quad (23)$$

(iv) For $k \in \mathbb{Z}$,

$$\varphi_j^{(n)}(k) = |\mathcal{D}(jk, 2n)|, \quad k \in \mathbb{Z}. \quad (24)$$

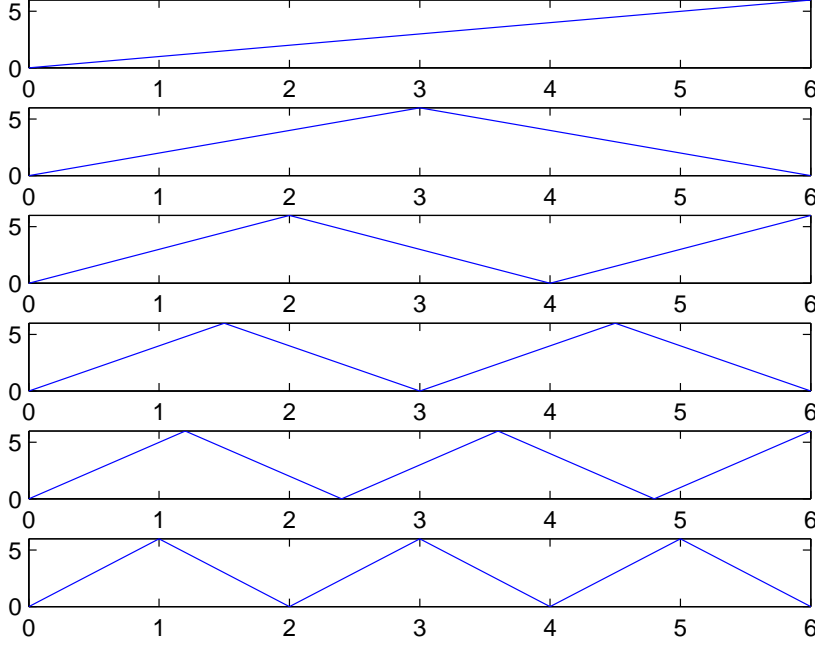


Figure 3: Functions $\varphi_j^{(6)}(x)$, $x \in [0, 6]$, for $j = 1, \dots, 6$.

Proof. Property (i) is obvious from construction. By (15), if $x \in \left[\frac{2k-1}{j}n, \frac{2k+1}{j}n\right]$ for $k \in \mathbb{Z}$, then

$$\varphi_j^{(n)}(x) = \begin{cases} 2kn - jx, & \text{if } x \in \left[\frac{2k-1}{j}n, \frac{2k}{j}n\right], \\ jx - 2kn, & \text{if } x \in \left[\frac{2k}{j}n, \frac{2k+1}{j}n\right]. \end{cases}$$

In particular, if $x \in \left[\frac{2k-1}{j}n, \frac{2k}{j}n\right]$, then $x + n/j \in \left[\frac{2k}{j}n, \frac{2k+1}{j}n\right]$. Thus,

$$\varphi_j^{(n)}(x + n/j) = j \left(x + \frac{n}{j}\right) - 2kn, \quad \varphi_j^{(n)}(x) = 2kn - jx,$$

and (22) follows. The case $x \in \left[\frac{2k}{j}n, \frac{2k+1}{j}n\right]$ is analyzed analogously.

Identity (23) is a straightforward consequence of the explicit formula (15).

Assume that $k \in \mathbb{Z}$ and $\mathcal{D}(jk, 2n) = r$; it means that there exists $t \in \mathbb{Z}$ such that $r = jk - 2nt \in [-n, n)$. Thus, $|\mathcal{D}(jk, 2n)| = |jk - 2nt|$. Furthermore, from the inequalities $-n \leq r < r$ it follows that

$$\frac{2t-1}{j}n \leq k < \frac{2t+1}{j}n.$$

Comparing it with the definition of $\varphi_j^{(n)}$ in (15), we establish (24). \square

Although function $\varphi_j^{(n)}$ is well defined on whole \mathbb{R} , we will be mainly interested in its values on the interval $[0, n]$. In particular, we need to study how $\varphi_j^{(n)}$ acts on integers $1, 2, \dots, n-1$:

Proposition 5 *Let $n, j \in \mathbb{N}$ with $\text{GCD}(j, n) = 1$.*

(i) *For any $j \in \mathbb{N}$, $\varphi_j^{(1)}(\mathbb{Z}) \subset \{0, 1\}$.*

(ii) *If $n > 1$ and j is odd, then*

$$\varphi_j^{(n)} : \{1, \dots, n-1\} \rightarrow \{1, \dots, n-1\}$$

is a bijection. In other words, $\varphi_j^{(n)}$ acts as a permutation on the set $\{1, \dots, n-1\}$.

(iii) *If $n > 1$ and j is even (and thus n is odd), then*

$$\varphi_j^{(n)} : \left\{1, \dots, \frac{n-1}{2}\right\} \rightarrow \left\{2m : m = 1, \dots, \frac{n-1}{2}\right\}$$

and

$$\varphi_j^{(n)} : \left\{\frac{n+1}{2}, \dots, n-1\right\} \rightarrow \left\{2m : m = 1, \dots, \frac{n-1}{2}\right\}$$

are bijections.

Proof. By construction, $\varphi_j^{(n)}(k) \in \{0, n\}$ if and only if $k = mn/j$, with $m \in \mathbb{Z}$. Since $\text{GCD}(j, n) = 1$, $mn/j \in \mathbb{Z}$ only if m is a multiple of j . Hence, $\varphi_j^{(n)}(k) \in \{0, n\}$ if and only if $k \in \{mn : m \in \mathbb{Z}\}$. With $n = 1$ this yields (i). Furthermore, if $n > 1$, then

$$\varphi_j^{(n)}(\{1, \dots, n-1\}) \subset \{1, \dots, n-1\}.$$

Hence, in order to prove (ii) it is sufficient to show that $\varphi_j^{(n)}$ is injective on $\{1, \dots, n-1\}$. Indeed, by (24), $\varphi_j^{(n)}(x) = r$ if $|\mathcal{D}(jx, 2n)| = r$, that is, if there exists $u \in \mathbb{Z}$ such that

$$jx \pm r = 2nu,$$

with an appropriate choice of the sign. Thus, if $\varphi_j^{(n)}(x) = \varphi_j^{(n)}(y)$, then there exists $u \in \mathbb{Z}$ such that

$$j(x \pm y) = 2nu,$$

again with an appropriate choice of the sign. However, since j is odd and $\text{GCD}(j, n) = 1$, we conclude that $\text{GCD}(j, 2n) = 1$. This means that $x \pm y$ must be divisible by $2n$. But $|x \pm y| < 2n$, so this identity is possible only if $x = y$.

If j is even, then a similar analysis shows that $k \rightarrow |\mathcal{D}(jk/2, n)|$ is injective both on

$$\left\{1, \dots, \frac{n-1}{2}\right\} \quad \text{and} \quad \left\{\frac{n+1}{2}, \dots, n-1\right\}.$$

It remains to use that by (24), $\varphi_j^{(n)}(k) = 2|\mathcal{D}(jk/2, n)|$, $k \in \mathbb{Z}$. This establishes (iii). \square

Proposition 6 *Let $m, n \in \mathbb{N}$, and $\text{GCD}(m, n) = d$. Then*

$$\text{GCD}(m, 2n) = d \quad \Leftrightarrow \quad m/d \text{ is odd,}$$

and

$$\text{GCD}(m, 2n) = 2d \quad \Leftrightarrow \quad m/d \text{ is even, and } n/d \text{ is odd.}$$

Proof. Assume that $d = 1$. It is obvious that $\text{GCD}(m, 2n) = 1$ only if m is odd, and viceversa, if m is odd and $\text{GCD}(m, n) = 1$, then necessarily $\text{GCD}(m, 2n) = 1$.

Analogously, if $\text{GCD}(m, 2n) = 2$, it means that m is even, and since m and n are coprime, n must be odd. The reciprocal is also trivially true: if m is even and $\text{GCD}(m, n) = 1$, then also $\text{GCD}(m/2, n) = 1$, and

$$\text{GCD}(m, 2n) = 2 \cdot \text{GCD}(m/2, n) = 2.$$

The general case is reduced to $d = 1$ by observing that

$$\text{GCD}(m, n) = d \quad \Leftrightarrow \quad \text{GCD}\left(\frac{m}{d}, \frac{n}{d}\right) = 1.$$

□

Now we are ready to prove the main result of this section.

Proof of the Main Lemma. Formula (17) is a straightforward consequence of (24). Since

$$\varphi_j^{(n)}\left(x + \frac{n}{d}\right) = \varphi_j^{(n)}\left(x + \frac{n}{j} \frac{j}{d}\right),$$

formula (16) follows from (22).

Since

$$\text{GCD}(j, n) = d \quad \Rightarrow \quad \text{GCD}\left(\frac{j}{d}, \frac{n}{d}\right) = 1,$$

we can apply Proposition 5 to function $\varphi_{j/d}^{(n/d)}$. In fact, statement (ii) is a straightforward consequence of (i) of Proposition 5 and formula (17).

Assume that $\text{GCD}(j, n) = \text{GCD}(j, 2n) = d < n$; by Proposition 6, j/d is odd. By (ii) of Proposition 5,

$$\varphi_{j/d}^{(n/d)} : \{1, \dots, n/d - 1\} \rightarrow \{1, \dots, n/d - 1\}$$

is a bijection, and by formula (17), this is valid also for

$$\varphi_j^{(n)} : \{1, \dots, n/d - 1\} \rightarrow \{dm : m = 1, \dots, n/d - 1\}.$$

Furthermore, any $k \in \{1, \dots, n - 1\}$ can be represented as

$$k = r + m \frac{n}{d}, \quad m \in \{0, 1, \dots, d - 1\}, \quad r \in \{0, 1, \dots, n/d - 1\}.$$

By (16),

$$\varphi_j^{(n)}(k) = \varphi_j^{(n)}\left(r + m \frac{n}{d}\right) = \begin{cases} n - \varphi_j^{(n)}(r), & \text{if } m \text{ is odd,} \\ \varphi_j^{(n)}(r), & \text{if } m \text{ is even.} \end{cases}$$

In consequence, for every $m \in \{0, 1, \dots, d-1\}$,

$$\varphi_j^{(n)} : \left\{ 1 + m \frac{n}{d}, 2 + m \frac{n}{d}, \dots, (m+1) \frac{n}{d} - 1 \right\} \rightarrow \{dm : m = 1, \dots, n/d-1\}$$

is a bijection. This proves (18)–(19).

On the other hand, if $\text{GCD}(j, 2n) = 2d < 2n$, then by Proposition 6, j/d is even and n/d is odd, and by (iii) of Proposition 5,

$$\varphi_{j/d}^{(n/d)} : \left\{ 1, \dots, \frac{n/d-1}{2} \right\} \rightarrow \left\{ 2m : m = 1, \dots, \frac{n/d-1}{2} \right\}$$

and

$$\varphi_{j/d}^{(n/d)} : \left\{ \frac{n/d+1}{2}, \dots, n/d-1 \right\} \rightarrow \left\{ 2m : m = 1, \dots, \frac{n/d-1}{2} \right\}$$

are bijections, so that by formula (17), this is valid also for

$$\varphi_j^{(n)} : \left\{ 1, \dots, \frac{n/d-1}{2} \right\} \rightarrow \left\{ 2dm : m = 1, \dots, \frac{n/d-1}{2} \right\}$$

and

$$\varphi_j^{(n)} : \left\{ \frac{n/d+1}{2}, \dots, n/d-1 \right\} \rightarrow \left\{ 2dm : m = 1, \dots, \frac{n/d-1}{2} \right\}.$$

Again, if $k \in \{1, \dots, n-1\}$, and

$$k = r + m \frac{n}{d}, \quad m \in \{0, 1, \dots, d-1\}, \quad r \in \{0, 1, \dots, n/d-1\},$$

we have by (16),

$$\varphi_j^{(n)}(k) = \varphi_j^{(n)}\left(r + m \frac{n}{d}\right) = \varphi_j^{(n)}(r).$$

In consequence, for every $m \in \{0, 1, \dots, d-1\}$,

$$\varphi_j^{(n)} : \left\{ 1 + m \frac{n}{d}, 2 + m \frac{n}{d}, \dots, \frac{n/d-1}{2} + m \frac{n}{d} \right\} \rightarrow \left\{ 2dm : m = 1, \dots, \frac{n/d-1}{2} \right\}$$

and

$$\varphi_j^{(n)} : \left\{ \frac{n/d+1}{2} + m \frac{n}{d}, \dots, \frac{n}{d} - 1 + m \frac{n}{d} \right\} \rightarrow \left\{ 2dm : m = 1, \dots, \frac{n/d-1}{2} \right\}$$

are bijections. This proves (20)–(21). \square

3 Discrete entropy for Chebyshev polynomials of the first kind

From the explicit formulas (13) for p_n is easy to compute that in this case

$$\ell_n^{-1}(\lambda) = n - \frac{1}{2} + \frac{1}{2} \frac{\sin(2n-1)\theta}{\sin \theta},$$

and for $\mathcal{S}_n(\lambda)$ defined in (10) we have

$$\begin{aligned}
\mathcal{S}_n(\lambda) &= -\log(\ell_n(\lambda)) - \ell_n(\lambda) \sum_{i=0}^{n-1} p_i^2(\lambda) \log(p_i^2(\lambda)) \\
&= -\log(\ell_n(\lambda)) - \ell_n(\lambda) \sum_{i=1}^{n-1} p_i^2(\lambda) \log(2 \cos^2(i\theta)(\lambda)) \\
&= -\log(\ell_n(\lambda)) - \ell_n(\lambda) \log(2) \sum_{i=1}^{n-1} p_i^2(\lambda) - \ell_n(\lambda) \sum_{i=1}^{n-1} p_i^2(\lambda) \log(\cos^2(i\theta)) \\
&= -\log(\ell_n(\lambda)) - \ell_n(\lambda) \log(2) (\ell_n^{-1}(\lambda) - 1) - 2\ell_n(\lambda) \sum_{i=1}^{n-1} \cos^2(i\theta) \log(\cos^2(i\theta)) ,
\end{aligned}$$

so that

$$\mathcal{S}_n(\lambda) = -\log(\ell_n(\lambda)) + \log(2) (\ell_n(\lambda) - 1) - 2\ell_n(\lambda) \sum_{i=1}^{n-1} \cos^2(i\theta) \log(\cos^2(i\theta)) . \quad (25)$$

Since for Chebyshev polynomials of the first kind and degree n the zeros are

$$\lambda_j^{(n)} = \cos\left(\frac{(2j-1)\pi}{2n}\right), \quad j = 1, \dots, n,$$

we see that $\ell_n(\lambda_j) = 1/n$ for $j = 1, \dots, n$. In particular, by (25), in this case

$$\mathcal{S}_{n,j} = \mathcal{S}_n(\lambda_j^{(n)}) = \log\left(\frac{n}{2}\right) + \frac{\log 2}{n} - \frac{2}{n} \widehat{\mathcal{S}}_{n,j}, \quad (26)$$

where

$$\widehat{\mathcal{S}}_{n,j} \stackrel{\text{def}}{=} \sum_{i=1}^{n-1} \cos^2\left(\frac{(2j-1)\pi}{2n} i\right) \log\left(\cos^2\left(\frac{(2j-1)\pi}{2n} i\right)\right). \quad (27)$$

Formulas (26)–(27) reduce the computation of $\mathcal{S}_{n,j}$ to the analysis of the *modified entropy* $\widehat{\mathcal{S}}_{n,j}$. But first we express $\widehat{\mathcal{S}}_{n,j}$ in terms of the auxiliary functions $\varphi_j^{(n)}$ defined by (15).

Proposition 7 For $j \in \mathbb{N}$,

$$\left| \cos\left(\frac{(2j-1)\pi}{2n} x\right) \right| = \left| \cos\left(\frac{\pi}{2n} \varphi_{2j-1}^{(n)}(x)\right) \right|, \quad x \in \mathbb{R}. \quad (28)$$

In particular,

$$\widehat{\mathcal{S}}_{n,j} = \sum_{k=1}^{n-1} \cos^2\left(\frac{\pi}{2n} \varphi_{2j-1}^{(n)}(k)\right) \log\left(\cos^2\left(\frac{\pi}{2n} \varphi_{2j-1}^{(n)}(k)\right)\right). \quad (29)$$

Proof. By (23), given $x \in \mathbb{R}$, either

$$\frac{\varphi_{2j-1}^{(n)}(x) - (2j-1)x}{2n} \in \mathbb{Z} \quad \text{or} \quad \frac{\varphi_{2j-1}^{(n)}(x) + (2j-1)x}{2n} \in \mathbb{Z}.$$

Hence, there exists $m \in \mathbb{Z}$ such that either

$$\frac{\pi}{2n} \varphi_j^{(n)}(x) = \frac{(2j-1)\pi}{2n} x + \pi m, \quad \text{or} \quad \frac{\pi}{2n} \varphi_j^{(n)}(x) = -\frac{(2j-1)\pi}{2n} x + \pi m,$$

and (28) follows. \square

Using the arithmetic properties of $\varphi_j^{(n)}$ established above, we can simplify the expression for the modified entropy:

Proposition 8 *Let $n \in \mathbb{N}$ and $j \in \{1, 2, \dots, n\}$. If $\text{GCD}(2j-1, n) = d$ then*

$$\widehat{S}_{n,j} = \begin{cases} 0 & \text{if } j = \frac{n+1}{2}, \\ d \sum_{k=1}^{(n/d)-1} \cos^2\left(\frac{\pi d}{2n} k\right) \log\left(\cos^2\left(\frac{\pi d}{2n} k\right)\right), & \text{otherwise.} \end{cases} \quad (30)$$

Furthermore,

$$\widehat{S}_{n,j} = \widehat{S}_{n,n-j+1}. \quad (31)$$

Proof. Observe that $d \leq n$, and for $j \in \{1, 2, \dots, n\}$ we have

$$d = n \quad \Leftrightarrow \quad j = \frac{n+1}{2}.$$

Furthermore, any term with index k in the sum (29), for which $\varphi_{2j-1}^{(n)}(k) \in \{0, n\}$, vanishes. Since $\text{GCD}(2j-1, n) = \text{GCD}(2j-1, 2n)$, formula (30) follows in a straightforward way from (29), (ii)–(iii) of the Main Lemma, and the commutativity of the sum.

Moreover,

$$\text{GCD}(2j-1, n) = \text{GCD}(2(n-j+1)-1, n). \quad (32)$$

Indeed, let $\text{GCD}(2j-1, n) = d$, so that $2j-1 = ds$, $n = dt$, where $s, t \in \mathbb{N}$ and $\text{GCD}(s, t) = 1$. By a well known characterization of coprime integers, there exists integers x, y such that $xs + yt = 1$. We have $2(n-j+1)-1 = (2t-s)d$, and

$$-x(2t-s) + (2x+y)t = xs + yt = 1 \quad \Rightarrow \quad \text{GCD}(2t-s, t) = 1,$$

so that $\text{GCD}(2(n-j+1)-1, n) = d$, which proves (32). In particular, $\widehat{S}_{n,j}$ and $\widehat{S}_{n,n-j+1}$ share the same value d in (30), which proves (31). \square

Next we find a series representation for $\widehat{S}_{n,j}$:

Proposition 9 *Let $n \in \mathbb{N}$, $j \in \{1, 2, \dots, n\}$, and $\text{GCD}(2j-1, n) = d$. Then*

$$\frac{1}{n} \widehat{S}_{n,j} = \frac{1}{2} (1 - 2 \log 2) + \sum_{s=1}^{\infty} \zeta(2s+1) \left(\frac{d}{2n}\right)^{2s+1}, \quad (33)$$

where $\zeta(\cdot)$ is the Riemann zeta function.

Remark 6 Observe that $d/2n \leq 1/2$, so that the series in the right hand side is convergent.

Proof. In order to find the value of $\widehat{S}_{n,j}$ we follow the strategy [3] of computing for $0 < 2 - \epsilon < q < 2 + \epsilon$ the l^q norms

$$N(q; h) \stackrel{\text{def}}{=} \sum_{j=0}^{n/d-1} \cos^q \left(\frac{\pi d}{2n} j \right) = \sum_{j=0}^{\pi/(2h)-1} \cos^q(hj), \quad h \stackrel{\text{def}}{=} \frac{\pi d}{2n}, \quad (34)$$

and considering the partial derivative of $N(q; h)$ with respect to q at $q = 2$:

$$\widehat{S}_{n,j} = 2d \sum_{j=0}^{\pi/(2h)-1} \cos^2(hj) \log \cos(hj) = 2d \left. \frac{\partial}{\partial q} N(q; h) \right|_{q=2}. \quad (35)$$

Observe that $N(q; h)$ are related to the Riemann sums of an integral:

$$hN(q; h) \simeq \int_0^{\pi/2} \cos^q(u) du = \int_0^{\pi/2} \left(\frac{\cos(u)}{\frac{\pi}{2} - u} \right)^q \left(\frac{\pi}{2} - u \right)^q du. \quad (36)$$

Consider $\cos^q(x)$ and $(\cos(x)/(\pi/2 - x))^q$ as analytic functions at $x = 0$ and $x = \pi/2$, respectively, whose single valued branches in the corresponding neighborhoods are fixed by

$$\cos^q(x)|_{x=0} = \left(\frac{\cos(x)}{\frac{\pi}{2} - x} \right)^q \Big|_{x=\pi/2} = 1.$$

Denote by

$$\cos^q(x) = 1 + \sum_{s=1}^{\infty} \alpha_s(q) x^s, \quad \left(\frac{\cos(x)}{\frac{\pi}{2} - x} \right)^q = 1 + \sum_{s=1}^{\infty} \beta_s(q) \left(\frac{\pi}{2} - x \right)^s$$

the Taylor expansions of these functions. Observe that $\alpha_s(q) = \beta_s(q) = 0$ for odd indices s . The Euler-Maclaurin summation formula for integrals with algebraic singularity at the end points (see [23]) yields

$$\begin{aligned} hN(q; h) &= h \sum_{j=0}^{\pi/(2h)-1} \cos^q(hj) = h + \int_0^{\pi/2} \cos^q(u) du + \\ &\quad + \sum_{s=0}^{\infty} \alpha_s(q) \zeta(-s) h^{s+1} + \sum_{s=0}^{\infty} \beta_s(q) \zeta(-s - q) h^{s+q+1}; \end{aligned} \quad (37)$$

at this stage we understand this identity in standard terms of an asymptotic expansion. Since $\zeta(0) = -1/2$ and $\zeta(-2j) = 0$ for $j \in \mathbb{N}$, formula in (37) reduces to

$$N(q; h) = \frac{1}{2} + \frac{1}{h} \int_0^{\pi/2} \cos^q(u) du + \sum_{s=0}^{\infty} \beta_{2s}(q) \zeta(-2s - q) h^{2s+q}. \quad (38)$$

Paper [23] addresses also the case of the Euler-Maclaurin summation formula providing full asymptotic expansion for integrands with logarithmic singularities at the end points. In fact,

these results from [23] can be obtained by formal differentiation of (38) with respect to q . Taking into account (35) we obtain:

$$\begin{aligned}\widehat{S}_{n,j} &= 2d \frac{\partial}{\partial q} N(q; h) \Big|_{q=2} \\ &= \frac{2d}{h} \int_0^{\pi/2} \cos^2(u) \log(\cos(u)) du - 2d \sum_{s=0}^{\infty} \beta_{2s}(2) \zeta'(-2(s+1)) h^{2(s+1)}.\end{aligned}\tag{39}$$

But

$$\int_0^{\pi/2} \cos^2(u) \log(\cos(u)) du = \frac{\pi}{8} (1 - \log 4),$$

and

$$\frac{\cos^2(z)}{\left(\frac{\pi}{2} - z\right)^2} = \sum_{s=0}^{\infty} \frac{(-1)^s 2^{2s+1}}{(2s+2)!} \left(\frac{\pi}{2} - z\right)^{2s},$$

so that

$$\beta_{2s}(2) = \frac{(-1)^s 2^{2s+1}}{(2s+2)!}.$$

Recalling the definition of h and gathering these formulas in (39), we obtain

$$\widehat{S}_{n,j} = \frac{n}{2} (1 - 2 \log 2) + d \sum_{s=1}^{\infty} (-1)^s \frac{\pi^{2s} \zeta'(-2s)}{(2s)!} \left(\frac{d}{n}\right)^{2s}.$$

Identity

$$\zeta'(-2s) = (-1)^s \zeta(2s+1) \frac{(2s)!}{\pi^{2s}} \frac{1}{2^{2s+1}}$$

yields now (33). It remains to observe that the series in the right hand side is convergent, thus this is a bona fide series expansion of $\widehat{S}_{n,j}$. \square

Corollary 10 *Let $n \in \mathbb{N}$, $j \in \{1, 2, \dots, n\}$, and $\text{GCD}(2j-1, n) = d$, then*

$$\frac{2}{n} \widehat{S}_{n,j} = 1 - 2 \log 2 - \mathcal{R} \left(\frac{d}{2n} \right),\tag{40}$$

with \mathcal{R} defined in (11).

Proof. It is an immediate consequence of (33) and (12). \square

Remark 7 It is easy to check that $\mathcal{R}(1/2) = 1 - \log(4)$, so that for $d = n$ we have $\widehat{S}_{n,j} = 0$ (cf. formula (30)).

It remains to use (40) in (26) in order to complete the proof of Theorem 1.

4 Entropy of Chebyshev polynomials of the second kind

From the explicit formulas (14) it follows that the zeros of the Chebyshev polynomials of the second kind of degree n are

$$\lambda_j^{(n)} = \cos\left(\frac{j\pi}{n+1}\right), \quad j = 1, \dots, n,$$

and

$$\ell_n^{-1}(\lambda) = \sum_{k=0}^{n-1} p_k^2(\lambda) = \frac{1}{2} \frac{n \sin(\theta) - \cos((n+1)\theta) \sin(n\theta)}{\sin^3(\theta)}, \quad \lambda = \cos \theta. \quad (41)$$

By (14), with $\lambda = \cos(\theta)$,

$$\begin{aligned} \sum_{k=1}^n p_{k-1}^2(\lambda) \log(p_{k-1}^2(\lambda)) &= \sum_{k=1}^n p_{k-1}^2(\lambda) \log(\sin^2(k\theta)) - \log(\sin^2(\theta)) \sum_{k=1}^n p_{k-1}^2(\lambda) \\ &= \sin^{-2}(\theta) \sum_{k=1}^n \sin^2(k\theta) \log(\sin^2(k\theta)) - \log(\sin^2(\theta)) \ell_n^{-1}(\lambda). \end{aligned}$$

Introducing the notation

$$\widehat{\mathcal{S}}_{n,j} \stackrel{\text{def}}{=} \sum_{k=1}^{n-1} \sin^2\left(\frac{kj\pi}{n}\right) \log\left(\sin^2\left(\frac{kj\pi}{n}\right)\right), \quad (42)$$

and using (9) we get

$$\begin{aligned} \mathcal{S}_{n,j} &= -\log(\ell_n(\lambda_j^{(n)})) - \ell_n(\lambda_j^{(n)}) \left(\frac{\widehat{\mathcal{S}}_{n+1,j}}{\sin^2(j\pi/(n+1))} - \frac{\log(\sin^2(j\pi/(n+1)))}{\ell_n(\lambda_j)} \right) \\ &= -\log(\ell_n(\lambda_j^{(n)})) + \log(\sin^2(j\pi/(n+1))) - \frac{\ell_n(\lambda_j^{(n)})}{\sin^2(j\pi/(n+1))} \widehat{\mathcal{S}}_{n+1,j}. \end{aligned} \quad (43)$$

Furthermore, by (41),

$$\ell_n^{-1}(\lambda_j^{(n)}) = \frac{n+1}{2 \sin^2\left(\frac{j\pi}{n+1}\right)}, \quad (44)$$

so that

$$\mathcal{S}_{n,j} = \mathcal{S}_n(\lambda_j^{(n)}) = \log\left(\frac{n+1}{2}\right) - \frac{2}{n+1} \widehat{\mathcal{S}}_{n+1,j}. \quad (45)$$

Again, formula (45) reduces the computation of $\mathcal{S}_{n,j}$ to the analysis of the *modified entropy* $\widehat{\mathcal{S}}_{n,j}$. We express $\widehat{\mathcal{S}}_{n,j}$ in terms of the auxiliary functions $\varphi_j^{(n)}$ defined by (15):

Proposition 11 For $j \in \mathbb{N}$,

$$\left| \sin\left(\frac{j\pi}{n} x\right) \right| = \left| \sin\left(\frac{\pi}{2n} \varphi_{2j}^{(n)}(x)\right) \right|, \quad x \in \mathbb{R}. \quad (46)$$

In particular,

$$\widehat{\mathcal{S}}_{n,j} = \sum_{k=1}^{n-1} \sin^2\left(\frac{\pi}{2n} \varphi_{2j}^{(n)}(k)\right) \log\left(\sin^2\left(\frac{\pi}{2n} \varphi_{2j}^{(n)}(k)\right)\right). \quad (47)$$

Proof. By (23), for every $x \in \mathbb{R}$ either

$$\frac{\varphi_{2j}^{(n)}(x) - 2jx}{2n} \in \mathbb{Z} \quad \text{or} \quad \frac{\varphi_{2j}^{(n)}(x) + 2jx}{2n} \in \mathbb{Z}.$$

Hence, given $x \in \mathbb{R}$ there exists $m \in \mathbb{Z}$ such that either

$$\frac{\pi}{2n} \varphi_{2j}^{(n)}(x) = \frac{j\pi}{n} x + \pi m, \quad \text{or} \quad \frac{\pi}{2n} \varphi_j^{(n)}(x) = -\frac{j\pi}{n} x + \pi m,$$

which implies (46). □

Proposition 12 *Let $n \in \mathbb{N}$ and $j \in \{1, 2, \dots, n-1\}$. Then*

$$\widehat{S}_{n,j} = \begin{cases} 0 & \text{if } j = n/2, \\ D \sum_{m=1}^{\frac{n}{D}-1} \sin^2\left(\frac{\pi d}{n} m\right) \log\left(\sin^2\left(\frac{\pi d}{n} m\right)\right), & \text{otherwise,} \end{cases} \quad (48)$$

where $D = \text{GCD}(2j, n)$ and $d = \text{GCD}(j, n)$.

Furthermore,

$$\widehat{S}_{n,j} = \widehat{S}_{n,n-j}. \quad (49)$$

Proof. Observe that $D \leq n$, and for $j \in \{1, 2, \dots, n-1\}$ we have

$$D = n \quad \Leftrightarrow \quad j = \frac{n}{2}.$$

It follows from (ii) of the Main Lemma (Section 2) that in this case $\varphi_{2j}^{(n)}(\mathbb{Z}) \in \{0, n\}$, so that we obtain formula (48) for $j = n/2$.

For $j \neq n/2$ we consider two cases. First, assume that $D = 2d$. Then,

$$\text{GCD}(2j, 2n) = 2 \cdot \text{GCD}(j, n) = 2d = D = \text{GCD}(2j, n).$$

Since $D < n$, by (18)–(19), $\varphi_{2j}^{(n)}(\{1, \dots, n-1\}) \setminus \{0, n\} = \{Dm : m = 1, \dots, \frac{n}{D} - 1\}$, and for any $m \in \{1, \dots, \frac{n}{D} - 1\}$, $\text{card}\{k \in \{1, \dots, n-1\} : \varphi_{2j}^{(n)}(k) = Dm\} = D$. Hence, (47) and the commutativity of the sum yield

$$\widehat{S}_{n,j} = D \sum_{m=1}^{\frac{n}{D}-1} \sin^2\left(\frac{\pi D}{2n} m\right) \log\left(\sin^2\left(\frac{\pi D}{2n} m\right)\right),$$

which proves (48) in this case.

Assume next that $D = d$. Then,

$$\text{GCD}(2j, 2n) = 2 \cdot \text{GCD}(j, n) = 2d = 2D = 2 \cdot \text{GCD}(2j, n).$$

Since $D < n$, by (20)–(21), $\varphi_{2j}^{(n)}(\{1, \dots, n-1\}) \setminus \{0, n\} = \{2Dm : m = 1, \dots, \frac{n/D-1}{2}\}$, and for any $m \in \{1, \dots, \frac{n-D}{2D}\}$, $\text{card}\{k \in \{1, \dots, n-1\} : \varphi_{2j}^{(n)}(k) = 2Dm\} = 2D$. Using (47) and the commutativity of the sum, we conclude that

$$\widehat{S}_{n,j} = 2D \sum_{m=1}^{\frac{n/D-1}{2}} \sin^2\left(\frac{\pi D}{n} m\right) \log\left(\sin^2\left(\frac{\pi D}{n} m\right)\right). \quad (50)$$

Furthermore,

$$\sin^2\left(\frac{\pi D}{n} m\right) = \sin^2\left(\pi - \frac{\pi D}{n} m\right) = \sin^2\left(\frac{\pi D}{n} \left(\frac{n}{D} - m\right)\right),$$

so that

$$\widehat{S}_{n,j} = 2D \sum_{m=1}^{\frac{n/D-1}{2}} \sin^2\left(\frac{\pi D}{n} \left(\frac{n}{D} - m\right)\right) \log\left(\sin^2\left(\frac{\pi D}{n} \left(\frac{n}{D} - m\right)\right)\right).$$

But

$$\left\{m = 1, \dots, \frac{n/D-1}{2}\right\} \cup \left\{\frac{n}{D} - m : m = 1, \dots, \frac{n/D-1}{2}\right\} = \left\{1, \dots, \frac{n}{D} - 1\right\}.$$

Thus,

$$\widehat{S}_{n,j} = D \sum_{m=1}^{\frac{n}{D}-1} \sin^2\left(\frac{\pi D}{n} m\right) \log\left(\sin^2\left(\frac{\pi D}{n} m\right)\right),$$

which concludes the proof of (48).

Finally, we have that $j = ds$, $n = dt$, where $s, t \in \mathbb{N}$ and $\text{GCD}(s, t) = 1$. Again by a characterization of coprime integers, there exists integers x, y such that $xs + yt = 1$. But $n - j = (t - s)d$, and

$$-x(t - s) + (x + y)t = xs + yt = 1 \quad \Rightarrow \quad \text{GCD}(t - s, t) = 1,$$

so that $\text{GCD}(n - j, n) = d$. Analogously, $\text{GCD}(2j, n) = D = \text{GCD}(2(n - j), n)$. Now (49) is a straightforward consequence of (48). \square

Next we find a series representation for $\widehat{S}_{n,j}$:

Proposition 13 *Let $n \in \mathbb{N}$, $j \in \{1, 2, \dots, n - 1\}$, and $\text{GCD}(j, n) = d$. Then*

$$\frac{2}{n} \widehat{S}_{n,j} = 1 - 2 \log 2 - \mathcal{R}\left(\frac{d}{n}\right), \quad (51)$$

where $\mathcal{R}(\cdot)$ has been introduced in (11).

Proof. Observe first that

$$\begin{aligned} & \sum_{m=1}^{n-1} \sin^2\left(\frac{\pi}{2n} m\right) \log\left(\sin^2\left(\frac{\pi}{2n} m\right)\right) \\ &= \sum_{m=1}^{n-1} \sin^2\left(\frac{\pi}{2n} (n - m)\right) \log\left(\sin^2\left(\frac{\pi}{2n} (n - m)\right)\right) \\ &= \sum_{m=1}^{n-1} \cos^2\left(\frac{\pi}{2n} m\right) \log\left(\cos^2\left(\frac{\pi}{2n} m\right)\right). \end{aligned} \quad (52)$$

Let $\text{GCD}(2j, n) = D$; assume first that $D = 2d$. By (48) and (52),

$$\begin{aligned}\widehat{S}_{n,j} &= D \sum_{m=1}^{\frac{n}{D}-1} \sin^2\left(\frac{\pi d}{n} m\right) \log\left(\sin^2\left(\frac{\pi d}{n} m\right)\right) \\ &= D \sum_{m=1}^{\frac{n}{D}-1} \sin^2\left(\frac{\pi D}{2n} m\right) \log\left(\sin^2\left(\frac{\pi D}{2n} m\right)\right) \\ &= D \sum_{m=1}^{\frac{n}{D}-1} \cos^2\left(\frac{\pi D}{2n} m\right) \log\left(\cos^2\left(\frac{\pi D}{2n} m\right)\right) \\ &= \frac{n}{2} \left(1 - 2 \log 2 - \mathcal{R}\left(\frac{D}{2n}\right)\right).\end{aligned}$$

where for the last identity we have used (30) and (40). Since $D = 2d$, this proves (51) in this case.

The remaining case is analyzed in a similar fashion. □

Using (51) in (45) we complete the proof of Theorem 2.

5 Further numerical experiments

In this section we present some results of numerical evaluation of the entropy $\mathcal{S}_{n,j}$ for several orthogonal polynomials. Computation has been carried out in Fortran 95, by complete diagonalization of the corresponding Jacobi matrix L_n in (4), using the routine STEVD of the LAPACK95 library [2, 6], which computes all the eigenvalues and eigenvectors of a given matrix by means of a divide and conquer algorithm [21].

As a first illustration we present the entropies $\mathcal{S}_{n,j}$, $n = 150, 151, 152$, for two values of the parameters α, β of Jacobi polynomials $P_n^{(\alpha,\beta)}$, given by the recurrence relation (3) with

$$b_i = \frac{2}{2i + \alpha + \beta} \sqrt{\frac{i(i + \alpha)(i + \beta)(i + \alpha + \beta)}{(2i + \alpha + \beta + 1)(2i + \alpha + \beta - 1)}}, \quad a_i = \frac{\alpha^2 - \beta^2}{(2i + \alpha + \beta)(2i + \alpha + \beta - 2)}.$$

In Figure 4 we can also observe the “peaks” explained for the Chebyshev polynomials, but unlike in the latter case, they are pointing both downwards and upwards. Furthermore, the value distribution close the endpoints of the interval clearly differs from the behavior in the bulk. The feature of the endpoint behavior is even more visible for the symmetric Pollaczek polynomials $p_n^\theta(\cdot; a)$ (Figure 5), given by the recurrence relation (3) with

$$b_i = \frac{1}{2} \sqrt{\frac{i(i + 2\theta - 1)}{(i + \theta + a)(i + \theta + a - 1)}}, \quad a_i = 0.$$

For $a = 0$ these polynomials reduce to the Jacobi (or more precisely, Gegenbauer) polynomials $P_n^{(\theta-1/2, \theta-1/2)}$. If $a > 0$, the orthogonality weight for the Pollaczek polynomials does not satisfy the Szegő condition due precisely to its exponentially fast decay at the end points of the interval $[-1, 1]$.

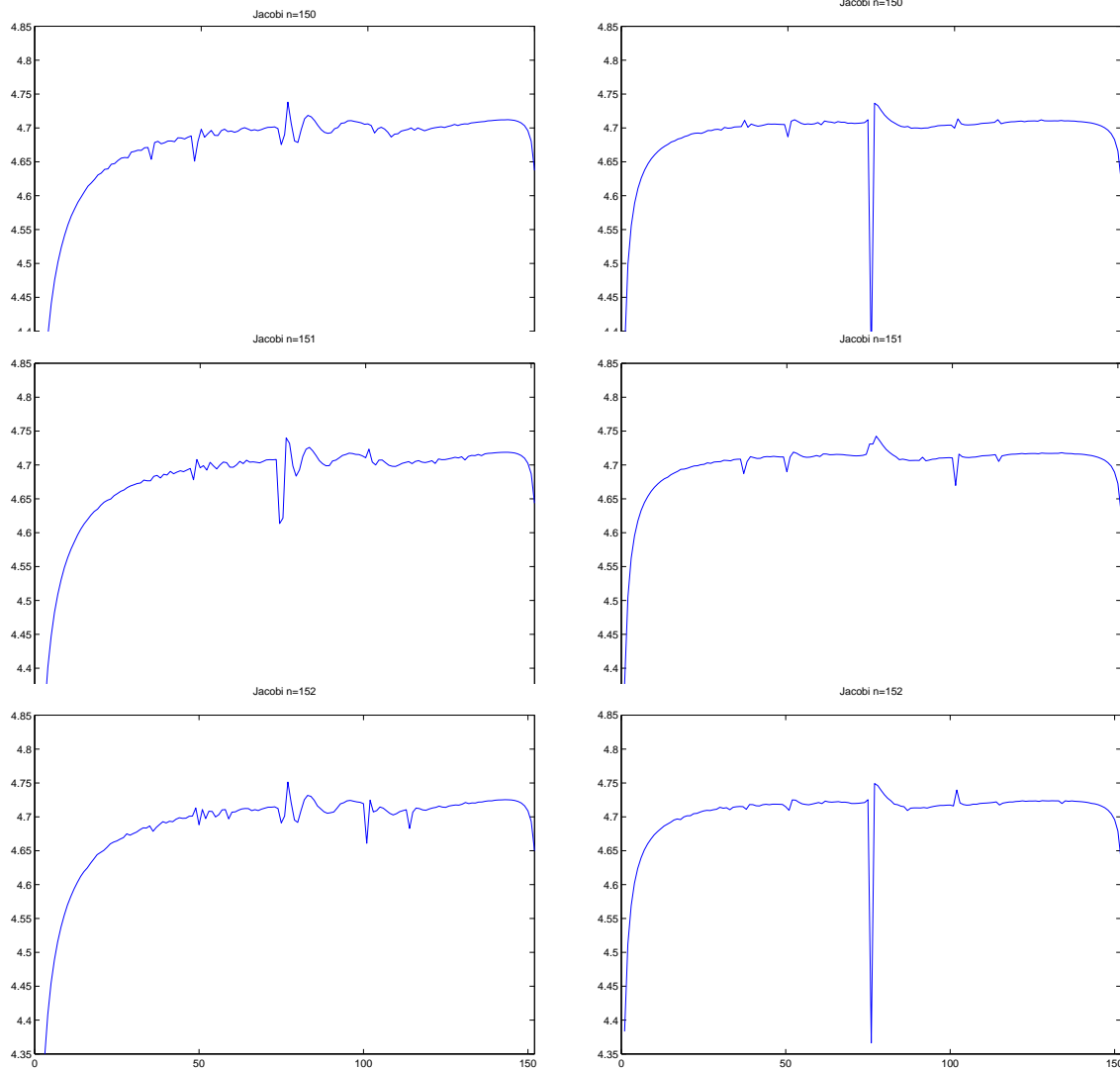


Figure 4: Jacobi polynomials $P_n^{(1.2,8.9)}$ (left) and $P_n^{(1.2,3.4)}$ (right): entropy $\mathcal{S}_{n,j}$ for $n = 150, 151$ and 152 .

A qualitatively different behavior is observed for the Meixner polynomials $M_n^{(\beta,c)}$, given by the recurrence relation (3) with

$$b_i = \frac{(ic(i + \beta - 1))^{1/2}}{1 - c}, \quad a_i = \frac{(i - 1)(1 + c) + c\beta}{1 - c}.$$

Recall that they are orthogonal with respect to the discrete measure (see e.g. [16, Chapter 6])

$$\mu = (1 - c)^\beta \sum_{k=0}^{\infty} \frac{(\beta)_k c^k}{k!} \delta_k.$$

From Figure 6 we observe that the value of the parameter c has greater impact on the behavior of the entropy $\mathcal{S}_{n,j}$ in comparison with the parameter β .

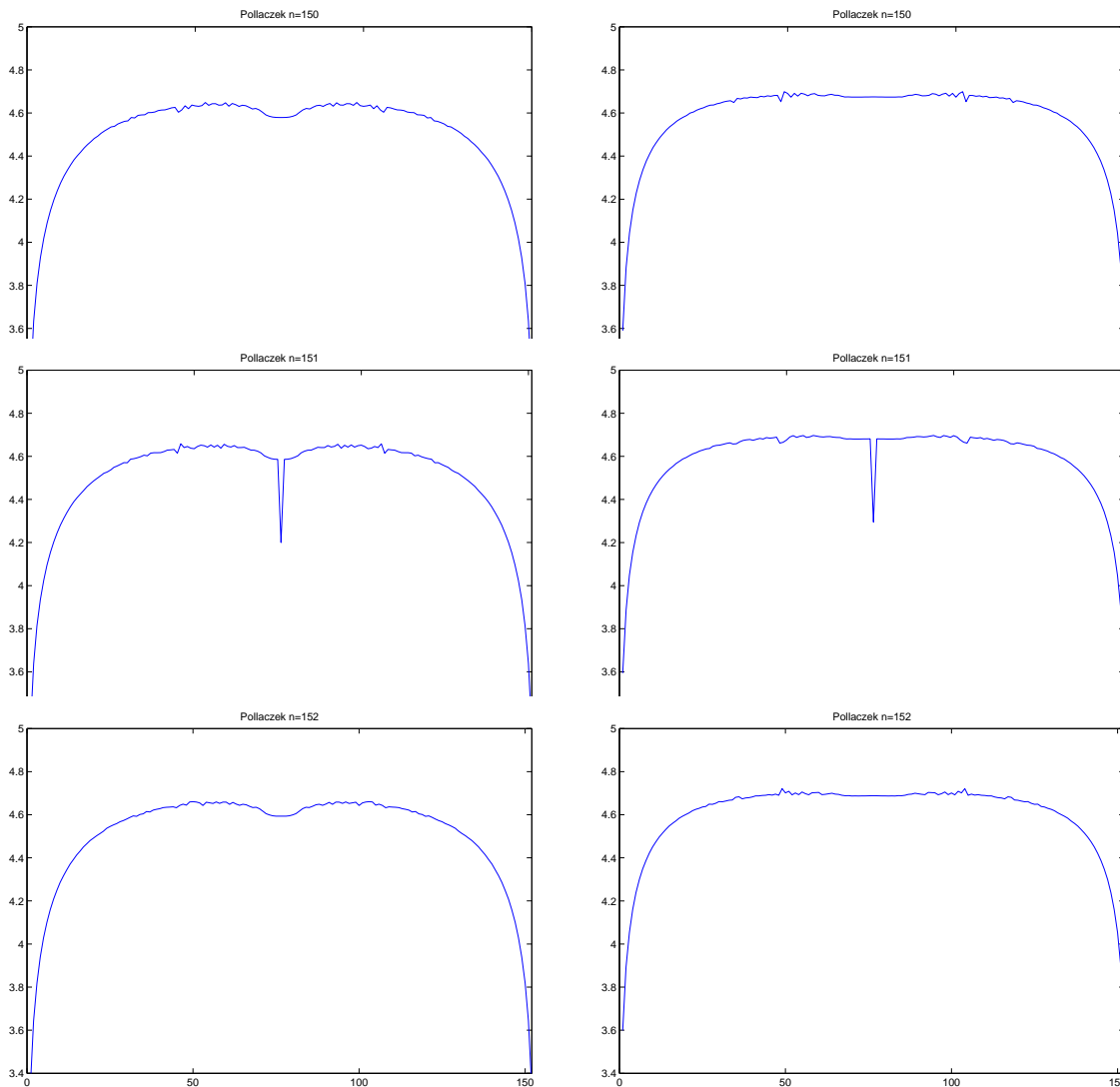


Figure 5: Pollaczek polynomials $p_n^{1,2}(x; 8.9)$ (left) and $p_n^{1,2}(x; 3.4)$ (right): entropy $\mathcal{S}_{n,j}$ for $n = 150, 151$ and 152 .

Finally, the evidence provided by all numerical experiments is sufficiently strong to conjecture that, after an appropriate rescaling and normalization, entropies $\mathcal{S}_{n,j}$ have a “semi-classical” limit as $n \rightarrow \infty$. The analysis of this asymptotic behavior is matter of a further research.

Acknowledgements

AIA was partially supported by Programm N1 of DMSRAS and grant RFBR 08-01-00179 of Russian Federation. JSD and RY were partially supported by Ministerio de Educación y Ciencia, grant FIS 2005-00973, and by Junta de Andalucía, grant FQM-207. AMF acknowledges support from Ministerio de Educación y Ciencia under grant MTM2005-09648-C02-01, and from Junta de Andalucía, grant FQM-229. Additionally, JSD, AMF and RY were partially

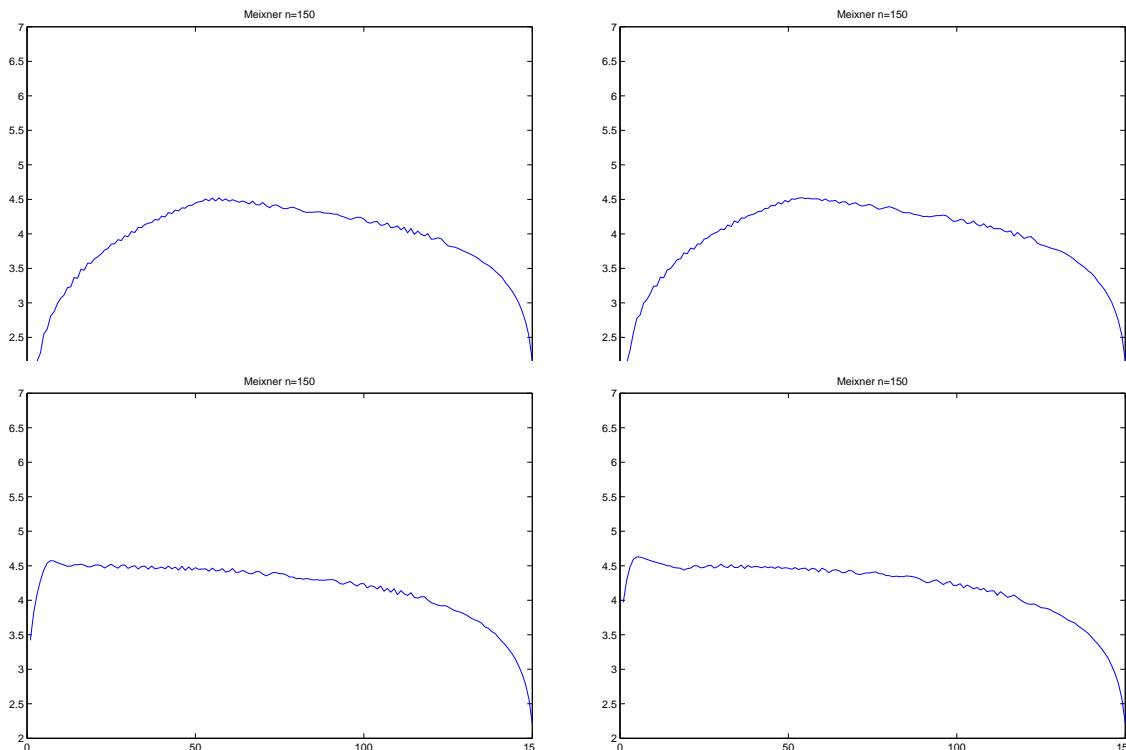


Figure 6: Entropy $\mathcal{S}_{150,j}$ for Meixner polynomials $M_n^{(\beta,c)}$ with $\beta = 3.4$, $c = 0.2$ (left top), $\beta = 8.9$, $c = 0.2$ (right top), $\beta = 3.4$, $c = 0.8$ (left bottom), $\beta = 8.9$, $c = 0.8$ (right bottom).

supported by the excellence grants FQM-481, and P06-FQM-01738 from Junta de Andalucía.

AIA also wishes to acknowledge the hospitality of the Universities of Almería and Granada, where this work was started.

References

- [1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover Publ., New York, 1972.
- [2] E. Anderson, Z. Bai, C. H. Bischof, S. Blackford, J. Demmel, J. J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. C. Sorensen. *LAPACK Users' Guide*. SIAM, Philadelphia, PA, USA, 3rd edition, 1999. <http://www.netlib.org/lapack/lug/>.
- [3] A. I. Aptekarev, V. S. Buyarov, and J. S. Dehesa. Asymptotic behavior of the L^p -norms and the entropy for general orthogonal polynomials. *Russian Acad. Sci. Sb. Math.*, 82(2):373–395, 1995.
- [4] A. I. Aptekarev, V. S. Buyarov, J. S. Dehesa, and W. Van Assche. Asymptotics for entropy integrals of orthogonal polynomials. *Russian Acad. Sci. Doklady Math.*, 53:47–49, 1996.

- [5] K. I. Babenko. An inequality in the theory of Fourier integrals. *Izv. Akad. Nauk SSSR Ser. Mat.*, 25:531–542, 1961. Engl. Transl. Am. Math. Soc., Transl., II. Ser. 44, 115–128 (1965).
- [6] V. A. Barker, S. Blackford, J. J. Dongarra, J. Du Croz, S. Hammarling, M. Marinova, J. Wa’sniewski, and P. Yalamov. *LAPACK95 Users’ Guide*. SIAM, Philadelphia, PA, USA, 2001. www.netlib.org/lapack95/lug95/.
- [7] B. Beckermann, A. Martínez-Finkelshtein, E. A. Rakhmanov, and F. Wielonsky. Asymptotic upper bounds for the entropy of orthogonal polynomials in the Szegő class. *J. Math. Physics*, 45(11):4239–4254, 2004.
- [8] I. Białynicki-Birula. Entropic uncertainty relations. *Phys. Lett.*, 103A:253–254, 1984.
- [9] I. Białynicki-Birula and J. Mycielsky. Uncertainty relations for information entropy in wave mechanics. *Commun. Math. Phys.*, 44:129–132, 1975.
- [10] V. Buyarov, J. S. Dehesa, A. Martínez-Finkelshtein, and J. Sánchez-Lara. Computation of the entropy of polynomials orthogonal on an interval. *SIAM J. Sci. Comp.*, 26(2):488–509, 2004.
- [11] V. S. Buyarov, J. S. Dehesa, A. Martínez-Finkelshtein, and E. B. Saff. Asymptotics of the information entropy for Jacobi and Laguerre polynomials with varying weights. *J. Approx. Theory*, 99(1):153–166, 1999.
- [12] V. S. Buyarov, P. López-Artés, A. Martínez-Finkelshtein, and W. Van Assche. Information entropy of Gegenbauer polynomials. *J. Phys. A*, 33(37):6549–6560, 2000.
- [13] J. S. Dehesa, A. Martínez-Finkelshtein, and J. Sánchez-Ruiz. Quantum information entropies and orthogonal polynomials. In *Proceedings of the Fifth International Symposium on Orthogonal Polynomials, Special Functions and their Applications (Patras, 1999)*, *J. Comput. Appl. Math.* 133: 23–46, 2001.
- [14] R.M. Dreizler and E.K.U. Gross. *Density Functional Theory: An Approach to the Quantum Mechanics*. Springer Verlag, Heidelberg, 1990.
- [15] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev. B*, 136:864–870, 1964.
- [16] M. E. H. Ismail. *Classical and quantum orthogonal polynomials in one variable*, volume 98 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005. With two chapters by Walter Van Assche, With a foreword by Richard A. Askey.
- [17] P. Jacquet and W. Szpankowski. Entropy computations via analytic de-Poissonization. *IEEE Trans. Inform. Theory*, 45(4):1072–1081, 1999.
- [18] C. Knessl. Integral representations and asymptotic expansions for Shannon and Rényi entropies. *Appl. Math. Lett.*, 11(2):69–74, 1998.
- [19] N.H. March. *Electron Density Theory of Atoms and Molecules*. Academic Press, New York, 1992.

- [20] A. Martínez-Finkelshtein and J. F. Sánchez-Lara. Shannon entropy of symmetric Polaczek polynomials. *J. Approx. Theory*, 145(1):55–80, 2007.
- [21] J. Rutter. A serial implementation of Cuppen’s divide and conquer algorithm for the symmetric eigenvalue problem. Technical Report CS-94-225, Department of Computer Science, University of Tennessee, Knoxville, TN, USA, March 1994. LAPACK Working Note 69.
- [22] M. Ohya and D. Petz. *Quantum Entropy and its Use*. Springer Verlag, New York, 1993.
- [23] A. Sidi. Euler-Maclaurin expansions for integrals with endpoint singularities: a new perspective. *Numer. Math.*, 98(2):371–387, 2004.

A.I. Aptekarev (aptekaa@spp.keldysh.ru)
 Keldysh Institute of Applied Mathematics,
 Miusskaya Pl. 4, 125047 Moscow, RUSSIA

J.S. Dehesa (dehesa@ugr.es)
 Institute Carlos I of Theoretical and Computational Physics,
 Granada University,
 Campus de Fuentenueva, 18071 Granada, SPAIN

A. Martínez-Finkelshtein (andrei@ual.es)
 Department of Statistics and Applied Mathematics, University of Almería,
 04120 Almería, SPAIN, and
 Institute Carlos I of Theoretical and Computational Physics,
 Granada University, SPAIN

R. Yáñez (ryanez@ugr.es)
 Institute Carlos I of Theoretical and Computational Physics,
 Granada University,
 Campus de Fuentenueva, 18071 Granada, SPAIN