COMPUTATION OF THE ENTROPY OF POLYNOMIALS ORTHOGONAL ON AN INTERVAL

V. BUYAROV*, J. S. DEHESA†, A. MARTíNEZ-FINKELSHTEIN‡, AND J. SÁNCHEZ-LARA§

Abstract. We give an effective method to compute the entropy for polynomials orthogonal on a segment of the real axis that uses as input data only the coefficients of the recurrence relation satisfied by these polynomials. This algorithm is based on a series expression for the mutual energy of two probability measures naturally connected with the polynomials. The particular case of Gegenbauer polynomials is analyzed in detail. These results are applied also to the computation of the entropy of spherical harmonics, important for the study of the entropic uncertainty relations as well as the spatial complexity of physical systems in central potentials.

Key words. Entropy, entropic uncertainty relation, orthogonal polynomials, Jacobi matrix, three term recurrence relation, Gegenbauer polynomials, spherical harmonics

AMS subject classifications. 33C45 33C55 33C90 33F05 42C05 65D20 65D30 81Q99

1. Introduction. The concept of information entropy in its continuous or discrete form has proved to be very fertile in numerous scientific branches because of its flexibility and multiple meanings [17, 19, 26, 38]. Indeed, it is used as a measure of disorder in thermodynamics [33], as a measure of uncertainty in statistical mechanics [21] as well as in classical and quantum information science [20, 25], as a measure of diversity in ecological structures, and as a criterion of classification of races and species in population dynamics [23], among others.

In quantum mechanics, the uncertainty in the localization of a particle in ordinary space is quantitatively measured by the so-called position information entropy [6]

$$S_\rho = -\int \rho(\vec{r}) \ln \rho(\vec{r}) d\vec{r}, \tag{1.1}$$

in a better and more convenient way than the Heisenberg's standard deviation of the quantum-mechanical probability density $\rho(\vec{r}) = |\psi(\vec{r})|^2$, where $\psi(\vec{r})$ is the wavefunction of its dynamical state. Similarly, the uncertainty in predicting the momentum of the particle is measured by the momentum information entropy S_γ of the density $\gamma(\vec{p}) = \left|\tilde{\psi}(\vec{p})\right|^2$, where the Fourier transform $\tilde{\psi}(\vec{p})$ of $\psi(\vec{r})$ is the wavefunction of the same state in the dual, conjugate or momentum space. These two quantities describe best [30, 31] the extent or spread of the position and momentum probability densities, respectively. Moreover, both entropies may decrease without bound when the corresponding density becomes more concentrated, i.e. when information in the associated space decreases. However, the entropy sum is bounded from below [5, 6]

$$S_\rho + S_\gamma \geq D(1 + \ln \pi),$$

where D is the dimensionality of the space (i.e. $D = 3$ for ordinary space). It expresses the impossibility to have a complete information of the position and momentum of the particle simultaneously. This is the so-called entropic uncertainty relation, which is a stronger version of the celebrated Heisenberg's uncertainty principle, a fundamental law of nature. This fact and the effective implementation of the density functional theory of complex many-electron systems [27], which uses the single-particle density as the basic variable, are responsible for the fact that the study of the entropy has become an ubiquitous tool in some areas (e.g. atomic and molecular physics, condensed matter theories). For instance, several maximum entropy methods

*Moscow State University (RUSSIA), email: BUYAROV@nw.math.msu.su

†Instituto Carlos I de Física Teórica y Computacional, Granada University (SPAIN), email: dehesa@ugr.es

‡University of Almería and Instituto Carlos I de Física Teórica y Computacional, Granada University (SPAIN), corresponding author, email: andrei@ual.es

§University of Almería (SPAIN), email: jlara@ual.es

based on the position and momentum entropies, as well as on their sum, have been developed [16, 21] and widely used [3, 15, 37] for determination of macroscopic quantities of natural systems. Nevertheless, the lack of a general theoretical methodology and of accurate numerical algorithms of computation these information entropies still prevents this approach from being more widely used.

The exact determination of the information entropies of complex many-particle systems is a formidable task. Only recently a small progress has been achieved using the theory of special functions, which in some cases allows to find closed formulas for the information entropies of the simplest 1-dimensional single-particle systems and the three-dimensional systems of particles moving in a central or spherically symmetric potential. For these systems the wavefunctions are controlled by some classical orthogonal polynomials (such as Gegenbauer, Laguerre or Hermite), and the determination of the corresponding information entropies boils down naturally to the computation of entropic functionals for sequences of orthogonal polynomials (cf. [3, 25, 36]; a state-of-the art of this topic up to 2001 is given in [11]).

Namely, for a positive unit Borel measure μ on $[-1, 1]$, let

$$p_n(x) = \gamma_n \prod_{j=1}^{n} \left(x - \zeta_j^{(n)} \right), \quad \gamma_n > 0, \tag{1.2}$$

be the corresponding *orthonormal* polynomials:

$$\int p_n(x)p_m(x)\, d\mu(x) = \delta_{mn}, \quad m, n \in \mathbb{Z}_+ .$$

We define the entropy of the polynomials $p_n(x)$ as

$$E_n = E_n(\mu) = -\int p_n^2(x) \ln\left(p_n^2(x) \right) d\mu(x) . \tag{1.3}$$

The presence of these integrals raises two questions. One is the study of their asymptotic behavior when $n \to \infty$, which has a special interest in the analysis of the highly-excited (Rydberg) states of numerous quantum-mechanical systems of hydrogenic-type [35]. In this sense there have been important contributions in the last few years [4, 8, 9, 13, 29]; for a detailed review, see [11]. A totally different problem is the explicit computation of (1.3) for every fixed n (up to a certain degree). Observe that a naive numerical evaluation of these functionals by means of quadratures is not convenient: since all the zeros of p_n belong to the interval of orthogonality, the increasing amount of integrable singularities spoils any attempt to achieve reasonable accuracy even for rather small n.

In this paper we present some theoretical results (Section 2), which allowed us to develop an algorithm for an effective and accurate numerical computation of the entropy of polynomials orthogonal on a segment of the real axis from the coefficients of the three-term recurrence relation which they satisfy (Section 3). In Section 4, we study in detail the case of Gegenbauer polynomials because of its own interest (as a very "representative" class of polynomials (1.2)) and because of their numerous applications; for instance, these polynomials control the angular component of the wavefunctions of single-particle systems in central potentials (cf. [11]). In Section 5, the results of several numerical experiments are discussed, illustrating both the accuracy and efficiency of the algorithm proposed here, and comparing it with other computing strategies used so far. Finally, the entropy of the spherical harmonics, which measures the spatial complexity of single-particle systems and physical systems with central potentials, is computed using the known relationship between spherical harmonics and Gegenbauer polynomials (Section 6).

2. Series representation of the entropy. The entropic functionals (1.3) can be restated in terms of the logarithmic potential theory. If μ and ν are Borel (generally speaking, real signed) measures on $\mathbb{C}$, we denote by

$$V(z; \mu) = -\int \ln |z - t|\, d\mu(t)$$

the logarithmic potential of μ, and define the following functionals:

$$I[\nu, \mu] = \int V(z; \nu) \, d\mu(z) = -\iint \ln|z - t| \, d\nu(t) \, d\mu(z),$$

is the mutual energy of μ and ν; $I(\mu) = I[\mu, \mu]$ is the logarithmic energy of μ, and

$$R[\nu, \mu] = -\int \ln\left(\frac{d\nu}{d\mu}\right) d\nu$$

is the relative entropy or the Kullback-Leibler information of ν and μ. From the Jensen inequality it immediately follows that if both μ and ν are positive unit measures, then $R[\nu, \mu] \leq 0$.

With the sequence of polynomials (1.2) we can associate naturally two sequences of probability measures on $[-1, 1]$:

$$\lambda_n = \frac{1}{n} \sum_{j=1}^{n} \delta_{\zeta_j^{(n)}} \qquad \text{and} \qquad d\nu_n(x) = p_n^2(x) \, d\mu(x). \tag{2.1}$$

Both measures are standard objects of study in the analytic theory of orthogonal polynomials. For instance, the normalized zero counting measure λ_n is closely connected with the n-th root asymptotics of p_n, and as was shown by Rakhmanov in his pioneering work [28], ν_n is associated with the behavior of the ratio p_{n+1}/p_n as $n \to \infty$.

With the notations introduced above the entropy (1.3) is equivalently rewritten as

$$E_n = R[\nu_n, \mu] \leq 0, \tag{2.2}$$

or as

$$E_n = -2\ln\gamma_n + 2\sum_{j=1}^{n} V(\zeta_j^{(n)}; \nu_n) = -2\ln\gamma_n + 2n \, I[\lambda_n, \nu_n]. \tag{2.3}$$

A standard procedure for computation of E_n employed so far (once the quadratures begin to fail), is based on formula (2.3). As it was shown in [12], the potential $V(x; \nu_n)$ oscillates on $[-1, 1]$ around the value $\ln 2$, which is the Robin (or extremal) constant of this interval. Moreover, the zeros $\zeta_j^{(n)}$ are points of local minima for the potential $V(x; \nu_n)$, hence in order to compute $E_n(\mu)$ we need to sum up the values of the logarithmic potential $V(z; \nu_n)$ at its local minima. Nevertheless, this procedure assumes an explicit computation of the zeros of p_n, with the drawback of the well-known instability and computational cost of this task as n grows large.

An exception in this sense is an algorithm, proposed in [9] for numerical computation of the entropy of Gegenbauer polynomials with integer parameters. It also involves finding zeros of certain polynomials generated recursively, but unlike in (2.3), the number of the zeros depends only on the parameter of the polynomial and not on its degree.

In this paper we propose a totally different approach to the computation of the entropic integrals (1.3), more in the spirit of standard numerical algorithms for orthogonal polynomials: it uses only the coefficients of the recurrence relation satisfied by these polynomials as the input data. It does not involve a solution of any nonlinear equation, and it can be carried out by performing matrix multiplication of structured (essentially, sparse) matrices. The algorithm is based on the formula contained in Theorem 2.1, which seems to be new.

We denote by L_μ^1 the class of measurable functions on $[-1, 1]$, absolutely integrable with respect to μ. Let $T_k(x) = \cos(k \arccos x)$ denote as usual the Chebyshev polynomials of the first kind. With the two measures introduced in (2.1) we define the double sequence of generalized moments

$$c_{k,n} = \int T_k(x) \, d\lambda_n(x), \quad m_{k,n} = \int T_k(x) \, d\nu_n(x), \quad k, n \geq 0. \tag{2.4}$$

Obviously, $|c_{k,n}| \leq 1$ and $|m_{k,n}| \leq 1$ for all values of k and n. One of the main results of this paper is the following

THEOREM 2.1. *Assume that μ is a unit Borel measure on $[-1,1]$. For $n \in \mathbb{N}$, let p_n be the corresponding orthonormal polynomial, and the unit measures λ_n and ν_n as defined in (2.1). Then for their mutual energy the following formula holds:*

$$I[\lambda_n, \nu_n] = \ln 2 + 2 \sum_{k=1}^{\infty} \frac{c_{k,n}\, m_{k,n}}{k} \,, \tag{2.5}$$

where the series in the right hand side is convergent.

Moreover, if we denote

$$M_n := \sup_{x \in [-1,1]} \int_{-1}^{1} \left| \frac{p_n^2(x) - p_n^2(t)}{x - t} \right| \, d\mu(x) < +\infty \,, \tag{2.6}$$

then, for $N \in \mathbb{N}$ we have

$$\left| I[\lambda_n, \nu_n] - \ln 2 - 2 \sum_{k=1}^{N} \frac{c_{k,n} m_{k,n}}{k} \right| \leq \frac{4M_n}{N+1} \,. \tag{2.7}$$

Proof. Following [12], we use the Fourier series of the logarithm [18, formula 1.514],

$$-\ln|1 - e^{i\varphi}| = \sum_{k=1}^{\infty} \frac{\cos k\varphi}{k} \,, \tag{2.8}$$

valid for almost all $\varphi \in [0, \pi]$ (see e.g. [10, Theorem 15.2]), which yields a representation of the logarithmic kernel

$$-\ln|x - t| = \ln 2 + 2 \sum_{k=1}^{\infty} \frac{1}{k} T_k(x) T_k(t) \,, \tag{2.9}$$

where for every $x \in [-1, 1]$ the series (in t) converges almost everywhere in $[-1, 1]$.

On the other hand, by the recurrence relation

$$T_{k+1}(x) = 2x\, T_k(x) - T_{k-1}(x) \,, \quad T_0(x) = 1 \,, \quad T_1(x) = x \,, \tag{2.10}$$

we have that

$$2(t - x)T_k(x)T_k(t) = q_k(x,t) - q_{k-1}(x,t) \,, \quad q_k(x,t) = T_{k+1}(t)T_k(x) - T_{k+1}(x)T_k(t) \,,$$

from where for $N \in \mathbb{N}$,

$$2(t - x) \sum_{k=1}^{N} \frac{T_k(x)T_k(t)}{k} = -(t - x) + \sum_{k=1}^{N-1} \frac{q_k(x,t)}{k(k+1)} + \frac{q_N(x,t)}{N} \,.$$

Hence,

$$\left| 2(t - x) \sum_{k=1}^{N} \frac{T_k(x)T_k(t)}{k} \right| \leq 4 \,, \quad t, x \in [-1, 1] \,.$$

Thus, if $f \in L_\mu^1$, we can apply Lebesgue dominated convergence theorem to (2.9) in order to assert that

$$-\int_{-1}^1 f(t)(t-x)\ln|t-x|\,d\mu(t) = \widehat{f}_0 \ln 2 + 2\sum_{k=1}^\infty \frac{\widehat{f}_k}{k} T_k(x)\,, \quad \widehat{f}_k = \int_{-1}^1 f(t)(t-x)T_k(t)\,d\mu(t)\,, \qquad (2.11)$$

and the series in the right hand side is convergent. Furthermore, we can estimate the remainder using that

$$2(t-x)\sum_{k=N+1}^\infty \frac{T_k(x)T_k(t)}{k} = -\frac{q_N(x,t)}{N+1} + \sum_{k=N+1}^\infty \frac{q_k(x,t)}{k(k+1)}\,,$$

from where

$$\left|2(t-x)\sum_{k=N+1}^\infty \frac{T_k(x)T_k(t)}{k}\right| \le \frac{4}{N+1}\,, \quad t,x \in [-1,1]\,,$$

so that

$$\left|2\sum_{k=N+1}^\infty \frac{\widehat{f}_k}{k} T_k(x)\right| \le \frac{4}{N+1}\int_{-1}^1 |f(t)|\,d\mu(t)\,. \qquad (2.12)$$

In particular, for $x \in [-1,1]$ we may take

$$f(t,\cdot) = \frac{p_n^2(t) - p_n^2(x)}{t-x}\,,$$

and formula (2.11) yields

$$-\int_{-1}^1 \left(p_n^2(t) - p_n^2(x)\right)\ln|t-x|\,d\mu(t) =$$

$$\ln 2 \int_{-1}^1 \left(p_n^2(t) - p_n^2(x)\right)d\mu(t) + 2\sum_{k=1}^\infty \frac{T_k(x)}{k}\int_{-1}^1 \left(p_n^2(t) - p_n^2(x)\right)T_k(t)\,d\mu(t)\,,$$

which using the definitions introduced above can be rewritten as

$$V(x;\nu_n) - p_n^2(x)V(x;\mu) = (1-p_n^2(x))\ln 2 + 2\sum_{k=1}^\infty \frac{T_k(x)}{k}\left(m_{k,n} - p_n^2(x)\int_{-1}^1 T_k(t)\,d\mu(t)\right)\,. \qquad (2.13)$$

Evaluating (2.13) at the zeros of p_n we obtain

$$V(\zeta_j^{(n)};\nu_n) = \ln 2 + 2\sum_{k=1}^\infty \frac{T_k(\zeta_j^{(n)})}{k} m_{k,n}\,, \quad j=1,2,\ldots,n\,.$$

Summing up these expressions for $j=1,2,\ldots,n$, we arrive at (2.5).

On the other hand, by (2.6) and (2.12),

$$\left|2\sum_{k=N+1}^\infty \frac{T_k(x)}{k}\int_{-1}^1 \left(p_n^2(t) - p_n^2(x)\right)T_k(t)\,d\mu(t)\right| \le \frac{4}{N+1}\int_{-1}^1 \left|\frac{p_n^2(t) - p_n^2(x)}{t-x}\right|\,d\mu(t) \le \frac{4M_n}{N+1}\,,$$

from where

$$\left|V(\zeta_j^{(n)};\nu_n) - \ln 2 - 2\sum_{k=1}^N \frac{T_k(\zeta_j^{(n)})}{k} m_{k,n}\right| = \left|2\sum_{k=N+1}^\infty \frac{T_k(\zeta_j^{(n)})}{k} m_{k,n}\right| \le \frac{4M_n}{N+1}\,, \quad j=1,2,\ldots,n\,,$$

and the estimate (2.7) follows. $\square$

COROLLARY 2.2. *With assumptions of Theorem 2.1,*

$$E_n = -2\ln\left(\frac{\gamma_n}{2^n}\right) + 4n\sum_{k=1}^{\infty}\frac{c_{k,n}m_{k,n}}{k}, \qquad (2.14)$$

and the error after truncating the series at the N-th term is bounded by the right hand side in (2.7).

REMARK: As our example of Gegenbauer polynomials in Section 4 shows, the bound (2.7) is usually too pessimistic.

3. Effective computation of E_n. Assume that we have as input data the coefficients of the three-term recurrence relation, satisfied by the orthonormal polynomials $p_n(x) = \gamma_n x^n + \ldots$,

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_n p_n(x) + a_n p_{n-1}(x), \qquad (3.1)$$

with $p_{-1} = 0$ and $p_0(x) = 1$. We form the infinite Jacobi matrix

$$J = \begin{pmatrix} b_0 & a_1 & 0 & 0 & \ldots \\ a_1 & b_1 & a_2 & 0 & \ldots \\ 0 & a_2 & b_2 & a_3 & \ldots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

and let $J_n = J(1:n, 1:n)$ denote its principal minor $n \times n$. Here and in what follows we occasionally use the MATLAB type notation to refer to elements of a matrix. Furthermore, e_n stands for the infinite (column) vector whose n-th element is 1 and the rest is 0, and $\langle a, b \rangle = a^H b$ is the standard scalar product in l^2 or $\mathbb{R}^k$ (the space where we multiply vectors is always clear from the context).

The following are very well known facts:

PROPOSITION 3.1. *Let p_n be the orthonormal polynomials (1.2) satisfying the recurrence relation (3.1). Then, with the notation above, for $n \geq 1$,*

i) the zeros $\zeta_j^{(n)}$, $j = 1, \ldots, n$, of p_n are eigenvalues of J_n, and $(p_0(\zeta_j^{(n)}), p_1(\zeta_j^{(n)}), \ldots, p_{n-1}(\zeta_j^{(n)})^T$ are corresponding eigenvectors. In particular, for $m \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$,

$$\sum_{j=1}^{n}\left[\zeta_j^{(n)}\right]^m = \text{trace of } [J_n]^m.$$

ii) If f is a polynomial then

$$\langle e_{n+1}, f(J)e_{n+1} \rangle = \int_{-1}^{1} f(x)p_n^2(x)w(x)\,dx. \qquad (3.2)$$

iii) For $r \geq n + 1$ and $m \in \mathbb{N}_0$,

$$\langle e_{n+1}, [J_r]^m e_{n+1} \rangle = \sum_{j=1}^{r}\Lambda_j^{(r)}\left[\zeta_j^{(r)}\right]^m p_n^2\left(\zeta_j^{(r)}\right), \qquad (3.3)$$

where $\Lambda_j^{(r)}$ are the Cotes-Christoffel numbers (Gauss quadrature weights) given by

$$\Lambda_j^{(r)} = \left[\sum_{i=0}^{r-1}p_i^2\left(\zeta_j^{(r)}\right)\right]^{-1}. \qquad (3.4)$$

iv) The leading coefficient of p_n satisfies $\gamma_n = (a_1 a_2 \ldots a_n)^{-1}$.

All these facts are classical (see e.g. [32]). Formula (3.2) can be found for instance in [34, §4.1.2], where it is proved for $f \in C[-1,1]$. Identity (3.3) is a straightforward consequence of (3.4), *i)*, and the spectral decomposition of the finite selfadjoint matrix J_r.

For practical computation of the left hand side in (3.2) we need the following result:

COROLLARY 3.2. *If $n \in \mathbb{N}_0$, $m, r \in \mathbb{N}$ and $r \geq n + (m+1)/2$, then the $(n+1, n+1)$ elements of J^m and $(J_r)^m$ coincide.*

Proof. It is well known that Gauss quadrature formula

$$\int_{-1}^{1} f(x) w(x) \, dx = \sum_{j=1}^{r} \Lambda_j^{(r)} f\left(\zeta_j^{(r)}\right)$$

is exact if the degree of the polynomial f is $\leq 2r - 1$. In particular,

$$\int_{-1}^{1} x^m p_n^2(x) w(x) \, dx = \sum_{j=1}^{r} \Lambda_j^{(r)} \left[\zeta_j^{(r)}\right]^m p_n^2\left(\zeta_j^{(r)}\right)$$

is exact if $m + 2n \leq 2r - 1$, and the statement follows from (3.2)–(3.3).[1] □

Assume that computing the entropy by means of formula (2.14) we decide to truncate the series therein at $k = N \in \mathbb{N}$. Then as a consequence of the previous corollary, in the right hand side of (3.2) we may use $T_k(J_r)$ instead of $T_k(J)$, with $r = n + 1 + [N/2]$. Moreover, in the right hand side of (3.2) it is sufficient to know only the $(n+1)$-th column of $f(J_r)$. Recalling that Chebyshev polynomials satisfy the recurrence relation (2.10), we can propose the following algorithm, where I_n stands for the $n \times n$ identity matrix:

ALGORITHM 1

(i)   `Compute` $\ln(\gamma_n/2^n) = -\ln\left[\prod_{j=1}^{n}(2a_j)\right]$ `recursively;`
        `choose` $N \in \mathbb{N}$ `at which truncate the series in (2.14);`
        `take` $T_0(J_n) = I_n$, $T_1(J_n) = J_n$, `and iterate`
                $T_k(J_n) = 2J_n T_{k-1}(J_n) - T_{k-2}(J_n)$, $k = 2, \ldots, N$,
        `computing`
(ii) $c_{k,n} = \operatorname{tr} T_k(J_n)/n$, $k = 1, \ldots, N$;
        `set` $r = n + 1 + [N/2]$; `starting with` $v_0 = I_r(:, n+1)$ `and` $v_1 = J_r(:, n+1)$,
        `iterate by the recurrence`
                $v_k = 2J_r v_{k-1} - v_{k-2}$, $k = 2, \ldots, N$,
        `computing`
(iii) $m_{k,n} = v_k(n+1)$, $k = 1, \ldots, N$;
        `substitute the results of (i)-(iii) in (2.14), terminating the series at` $k = N$.

Observe that this algorithm starts from the spectral data as the only input, and performs multiplication of structured matrices, without solving any kind of equation. This can be efficiently implemented, for instance, using the known algorithms for sparse matrix multiplication.

On the other hand, in order to satisfy conditions of Corollary 3.2 it is necessary to know in advance the truncation term N for the series in (2.14), for which we need an a priori bound on the error. The bound in (2.7) can be used, but usually it overestimates the error yielding values of N much larger than needed. In Section 4, we discuss the selection of the truncation term in the particular case of Gegenbauer polynomials.

---

[1] This short and elegant proof was suggested by one of the anonymous referees whose contribution we gratefully acknowledge.

4. Entropy computation for Gegenbauer polynomials. In this Section we test our approach on the computation of the entropy of the Gegenbauer polynomials.

For $\lambda > -1/2$ let

$$c_\lambda = \frac{\Gamma(\lambda + 1)}{\sqrt{\pi}\,\Gamma(\lambda + 1/2)}\,. \tag{4.1}$$

It is easy to verify that $w^\lambda(x) = c_\lambda(1 - x^2)^{\lambda - 1/2}$ is a positive unit weight on $[-1, 1]$. Let C_k^λ denote the Gegenbauer polynomial of degree k and parameter λ, orthogonal with respect to w^λ on this interval, and normalized by the value at $x = 1$,

$$C_k^\lambda(1) = \binom{k + 2\lambda - 1}{k}; \tag{4.2}$$

this is a standard normalization, adopted for instance in [1] and in [32]. Straightforward computation shows that

$$G_k^\lambda(x) = \left(\frac{k!\,(k + \lambda)\Gamma(2\lambda)}{\lambda\Gamma(k + 2\lambda)}\right)^{1/2} C_k^\lambda(x) = \gamma_k^\lambda\,x^k + \text{lower degree terms} \tag{4.3}$$

are the Gegenbauer polynomials orthonormal with respect to $w^\lambda(x)$.

In this Section we will use the superscript λ in the previously introduced notation when we want to make the dependence on the parameter λ explicit; for instance,

$$E_n^\lambda = -\int_{-1}^1 \left(G_n^\lambda(x)\right)^2 \ln\left(G_n^\lambda(x)\right)^2 w^\lambda(x)\,dx\,. \tag{4.4}$$

For the time being, only few explicit formulas for the entropy of orthonormal Gegenbauer polynomials are known. Namely, for $\lambda = 0$ and $\lambda = 1$ (Chebyshev polynomials of the first and second kind, respectively) it is not difficult to prove (cf. [12, 35]) that

$$E_n^0 = \log 2 - 1, \qquad E_n^1 = -\frac{n}{n + 1}\,. \tag{4.5}$$

Furthermore, case $\lambda = 2$ was studied in [7] and [9], establishing that

$$E_n^2 = \log\left(\frac{n + 3}{3(n + 1)}\right) - \frac{n^3 - 5n^2 - 29n - 27}{(n + 1)(n + 2)(n + 3)} - \frac{1}{n + 2}\left(\frac{n + 3}{n + 1}\right)^{n + 2}. \tag{4.6}$$

This formula does not allow to expect a short and elegant expression for all E_n^λ, even for integer values of λ.

For $\lambda \in \mathbb{N}$, $\lambda \geq 2$, an alternative algorithm has been proposed in [9]; it expresses the entropy in terms of the zeros of certain polynomials generated recursively, whose number, $2\lambda - 2$, depends only on the parameter of the polynomial and not on its degree.

This approach (valid only for integer values of λ), is more efficient than the direct computation of E_n^λ by quadrature when n grows large, and allows also to find constructively the first terms of the asymptotic expansion of the entropy when $n \to \infty$ and λ is fixed (cf. [4, 9, 29]):

$$E_n^\lambda = \mathcal{E}_0^\lambda + \frac{\mathcal{E}_1^\lambda}{n} + \dots, \quad \mathcal{E}_0^\lambda = -1 - \ln\frac{\Gamma(2\lambda)}{\Gamma(\lambda)\Gamma(\lambda + 1)}\,. \tag{4.7}$$

In this Section we will apply the general approach, described in Section 3, to the efficient computation of the entropy E_n^λ. It is well known that the polynomials G_n^λ satisfy the three-term recurrence relation

$$x\,G_n^\lambda(x) = a_{n+1}G_{n+1}^\lambda(x) + a_n G_{n-1}^\lambda(x),$$

where

$$a_n = \frac{1}{2} \left[\frac{n(n + 2\lambda - 1)}{(n + \lambda - 1)(n + \lambda)} \right]^{1/2} > 0 \, .$$

In particular, since the leading coefficient $\gamma_n^\lambda = (a_1 a_2 \ldots a_n)^{-1}$, we have

$$-2 \ln \left(\gamma_n^\lambda / 2^n \right) = \ln \left[\prod_{j=1}^n \frac{j(j + 2\lambda - 1)}{(j + \lambda - 1)(j + \lambda)} \right] \, . \tag{4.8}$$

Consider the values defined in (2.4). By symmetry,

$$c_{k,n}^\lambda = m_{k,n}^\lambda = 0 \text{ for } k \text{ odd.} \tag{4.9}$$

Observe that for $\lambda > -1/2$ the orthogonality weight satisfies the assumptions of Theorem 2.1. Thus, taking into account (4.8), formula (2.14) for the Gegenbauer polynomials reads as

$$E_n^\lambda = \ln \left(\prod_{j=1}^n \frac{j(j + 2\lambda - 1)}{(j + \lambda - 1)(j + \lambda)} \right) + 2n \sum_{k=1}^\infty \frac{c_{2k,n}^\lambda m_{2k,n}^\lambda}{k} \, , \quad \lambda > -1/2 \, , \quad n \in \mathbb{N} \, . \tag{4.10}$$

PROPOSITION 4.1. *For $\lambda \in \mathbb{N}_0$, the series in (4.10) is terminating at $k = n + \lambda$.*

Proof. This is a straightforward consequence of the fact that in this case $w^\lambda(x)(1 - x^2)^{1/2}$ is a polynomial of degree λ, and for $k > 2n + 2\lambda$ we can use orthogonality of T_k in the definition of $m_{k,n}$ in (2.4). $\square$

Taking advantage of (4.9), we can use the recurrence formula for even Chebyshev polynomials: $T_0(x) = 1$, $T_2(x) = 2x^2 - 1$, and

$$T_{2k}(x) = 2\, T_2(x)\, T_{2k-2}(x) - T_{2k-4}(x) \, , \quad k \geq 2 \, . \tag{4.11}$$

Then Algorithm 1 described in Section 3 takes the following form for the Gegenbauer polynomials:

ALGORITHM 2

(i)   Find $-2 \ln \left(\gamma_n^\lambda / 2^n \right)$ by (4.8) recursively.
Choose a value $N \geq n + \lambda$ where to truncate the series in (4.10),
   unless $\lambda \in \mathbb{N}_0$; in such a case, $N = n + \lambda$.
Take $T_0(J_n) = I_n$, $\widehat{J}_n = T_2(J_n) = 2J_n^2 - I_n$, and iterate
   $$T_{2k}(J_n) = 2\widehat{J}_n \, T_{2k-2}(J_n) - T_{2k-4}(J_n), \ \ k = 2, \ldots, N \, ,$$
computing
(ii) $c_{2k,n}^\lambda = \mathrm{tr}\, T_{2k}(J_n)/n, \ \ k = 1, \ldots, N$.
Set $r = n + 1 + [N/2]$ and $\widehat{J}_r = T_2(J_r) = 2J_r^2 - I_r$.
Starting with $v_0 = I_r(:, n+1)$ and $v_2 = \widehat{J}_r(:, n+1)$,
   iterate by the recurrence
   $$v_{2k} = 2\widehat{J}_r \, v_{2k-2} - v_{2k-4}, \ \ k = 2, \ldots, N \, ,$$
computing
(iii) $m_{2k,n}^\lambda = v_{2k}(n+1), \ \ k = 1, \ldots, N$.
Substitute the results of (i)-(iii) in (4.10),
   terminating the series at $k = N$.

In the implementation it is convenient to use subroutines for sparse matrix multiplication; as it was mentioned, the coefficient in (4.8) is better to compute recursively, avoiding possible overflows.

In order to obtain an a priori bound for the error (and thus to find the truncation term N) we can use the explicit expression for the coefficients $m_{2k,n}^{\lambda}$. In [35], $m_{2k,n}^{\lambda}$ were expressed in terms of Wilson polynomials of degree n and parameters depending on λ and k. Nevertheless, the expression which appears there has indeterminacies for integer values of λ, which is inconvenient for evaluation. We are interested in an alternative formula for $m_{2k,n}^{\lambda}$.

When a linearization formula of the type

$$T_k(x)\, p_n(x) = \sum_{j=0}^{n+k} \ell_{j,k,n} p_j(x)\,,$$

is available the coefficients $m_{k,n}$ in (2.4) can be found observing that $m_{k,n} = \ell_{n,k,n}$. Nevertheless, we were unable to find the explicit expression in the literature, and we establish a formula for $m_{2k,n}^{\lambda}$ based on the hypergeometric representation for the Chebyshev and Gegenbauer polynomials.

THEOREM 4.2. *For the orthonormal Gegenbauer polynomials G_n^{λ} the coefficients $m_{2k,n}^{\lambda}$ defined in (2.4) satisfy*

$$m_{2k,n}^{\lambda} = \frac{n+\lambda}{n!} \sum_{j=0}^{n} (-1)^j \binom{n}{j} \frac{(2\lambda+j)_n}{j+\lambda} \frac{(-j-\lambda)_k}{(j+\lambda+1)_k}\,, \quad k \geq 1\,. \tag{4.12}$$

Alternatively, for $k > n + \lambda$,

$$m_{2k,n}^{\lambda} = -\sin(\pi\lambda) \frac{n+\lambda}{\pi\, n!} \sum_{j=0}^{n} \binom{n}{j} \frac{(2\lambda+j)_n}{j+\lambda} \Gamma^2(j+\lambda+1) \frac{\Gamma(k-j-\lambda)}{\Gamma(k+j+\lambda+1)}\,. \tag{4.13}$$

In particular, for $\lambda \in \mathbb{N}_0$, $m_{2k,n} = 0$ for all $k \geq n + \lambda + 1$. Moreover, for any $\lambda > -1/2$ and $k > n + \lambda$, we have $|m_{2k+2,n}| \leq |m_{2k,n}|$, and

$$|m_{2k,n}| = O\left(\frac{1}{k^{2\lambda+1}}\right)\,, \quad k \to +\infty\,. \tag{4.14}$$

Here $(a)_k = \Gamma(a+k)/\Gamma(a)$ denotes the Pochhammer's symbol.

Proof. The following hypergeometric representation for the Chebyshev and Gegenbauer polynomials is well known (see e.g. [22, formulas 8.942 and 8.932]):

$$T_{2k}(x) = {}_2F_1\left(\begin{array}{c} -2k,\, 2k \\ 1/2 \end{array}\middle|\, \frac{1-x}{2}\right), \quad C_n^{\lambda}(x) = \frac{(2\lambda)_n}{n!}\, {}_2F_1\left(\begin{array}{c} -n,\, n+2\lambda \\ \lambda+1/2 \end{array}\middle|\, \frac{1-x}{2}\right).$$

The quadratic transformation

$${}_2F_1\left(\begin{array}{c} -2n,\, 2n+2\alpha+1 \\ \alpha+1 \end{array}\middle|\, \frac{1-x}{2}\right) = {}_2F_1\left(\begin{array}{c} -n,\, n+\alpha+1/2 \\ \alpha+1 \end{array}\middle|\, 1-x^2\right)$$

yields

$$T_{2k}(x) = {}_2F_1\left(\begin{array}{c} -k,\, k \\ 1/2 \end{array}\middle|\, 1-x^2\right), \quad C_n^{\lambda}(x) = \frac{(2\lambda)_n}{n!}\, {}_2F_1\left(\begin{array}{c} -n/2,\, n/2+\lambda \\ \lambda+1/2 \end{array}\middle|\, 1-x^2\right),$$

and Clausen's identity (cf. [2, p. 116, problem 13] or [14, Ch. IV, section 4.3])

$$\left({}_2F_1\left(\begin{array}{c} a,\, b \\ a+b+1/2 \end{array}\middle|\, x\right)\right)^2 = {}_3F_2\left(\begin{array}{c} 2a,\, 2b,\, a+b \\ 2a+2b,\, a+b+1/2 \end{array}\middle|\, x\right)$$

gives the following representation:

$$[C_n^\lambda(x)]^2 = \left(\frac{(2\lambda)_n}{n!}\right)^2 \, {}_3F_2\left(\begin{array}{c} -n, \, n+2\lambda, \, \lambda \\ 2\lambda, \, \lambda+1/2 \end{array} \middle| 1-x^2\right).$$

We use these formulas in order to compute the integral

$$D_{2k,n} = \int_{-1}^1 T_{2k}(x) \, C_n^\lambda(x)^2 \, (1-x^2)^{\lambda-1/2}dx = \left(\frac{(2\lambda)_n}{n!}\right)^2$$

$$\times \sum_{i=0}^k \frac{(-k)_i \, (k)_i}{(1/2)_i \, i!} \left(\sum_{j=0}^n \frac{(-n)_j \, (n+2\lambda)_j \, (\lambda)_j}{(2\lambda)_j(\lambda+1/2)_j \, j!} \int_{-1}^1 (1-x^2)^{i+j+\lambda-1/2}dx\right)$$

$$= \sqrt{\pi} \left(\frac{(2\lambda)_n}{n!}\right)^2 \sum_{i=0}^k \sum_{j=0}^n \frac{(-k)_i \, (k)_i}{(1/2)_i \, i!} \frac{(-n)_j \, (n+2\lambda)_j \, (\lambda)_j}{(2\lambda)_j \, (\lambda+1/2)_j \, j!} \frac{\Gamma(i+j+\lambda+1/2)}{\Gamma(i+j+\lambda+1)}.$$

Interchanging the order of summation we get

$$D_{2k,n} = \sqrt{\pi} \left(\frac{(2\lambda)_n}{n!}\right)^2 \sum_{j=0}^n {}_3F_2\left(\begin{array}{c} -k, \, k, \, j+\lambda+1/2 \\ 1/2, \, j+\lambda+1 \end{array} \middle| 1\right) \frac{(-n)_j \, (n+2\lambda)_j \, (\lambda)_j}{(2\lambda)_j \, (\lambda+1/2)_j \, j!} \frac{\Gamma(j+\lambda+1/2)}{\Gamma(j+\lambda+1)}.$$

Using the Pfaff-Saalschutz identity,

$$ {}_3F_2\left(\begin{array}{c} -k, \, a, \, b \\ c, 1+a+b-c-k \end{array} \middle| 1\right) = \frac{(c-a)_k \, (c-b)_k}{(c)_k \, (c-a-b)_k},$$

the integral $D_{2k,n}$ becomes

$$D_{2k,n} = \sqrt{\pi} \left(\frac{(2\lambda)_n}{n!}\right)^2 \sum_{j=0}^n \frac{(1/2-k)_k \, (-j-\lambda)_k}{(1/2)_k \, (-k-j-\lambda)_k} \frac{(-n)_j \, (n+2\lambda)_j \, (\lambda)_j}{(2\lambda)_j \, (\lambda+1/2)_j \, j!} \frac{\Gamma(j+\lambda+1/2)}{\Gamma(j+\lambda+1)}.$$

Taking into account the normalization factors in (4.1) and (4.3) we obtain

$$m_{2k,n}^\lambda = \frac{n! \, (n+\lambda)}{\lambda \, (2\lambda)_n} \frac{\Gamma(\lambda+1)}{\sqrt{\pi} \, \Gamma(\lambda+1/2)} \int_{-1}^1 T_{2k}(x) C_n^\lambda(x)^2 (1-x^2)^{\lambda-1/2}dx$$

$$= \frac{(n+\lambda) \, \Gamma(\lambda) \, (2\lambda)_n}{n! \, \Gamma(\lambda+1/2)} \sum_{j=0}^n \frac{(1/2-k)_k \, (-j-\lambda)_k}{(1/2)_k \, (-k-j-\lambda)_k} \frac{(-n)_j \, (n+2\lambda)_j \, (\lambda)_j}{(2\lambda)_j \, (\lambda+1/2)_j \, j!} \frac{\Gamma(j+\lambda+1/2)}{\Gamma(j+\lambda+1)}$$

$$= (-1)^k \frac{n+\lambda}{n!} \sum_{j=0}^n (-1)^j \left(\begin{array}{c} n \\ j \end{array}\right) \frac{(-j-\lambda)_k}{(-k-j-\lambda)_k} \, (2\lambda+j)_n \, \frac{1}{j+\lambda}$$

$$= \frac{n+\lambda}{n!} \sum_{j=0}^n (-1)^j \left(\begin{array}{c} n \\ j \end{array}\right) \frac{(2\lambda+j)_n}{j+\lambda} \frac{(-j-\lambda)_k}{(j+\lambda+1)_k},$$

where we have used standard properties of the Gamma function and Pochhammer's symbol. This proves (4.12). Formula (4.13) follows easily from (4.12) and the well known relation $\Gamma(x)\Gamma(1-x) = \pi/\sin(\pi x)$ applied with $x = j+\lambda+1 > 0$. In particular, it shows that $|m_{2k+2,n}| \leq |m_{2k,n}|$; finally, (4.14) is a consequence of (4.13) and Stirling formula. $\square$

In order to discuss the truncation error in (4.10) we need to introduce the following notation: fixed n, $\lambda > -1/2$, $\lambda \notin \mathbb{N}_0$, and $N \in \mathbb{N}$, $N > n+\lambda$, let

$$R_n^\lambda(N) = 2n \sum_{k=N}^\infty \frac{c_{2k,n}^\lambda \, m_{2k,n}^\lambda}{k};$$

$|R_n^\lambda(N)|$ is the absolute error of approximation of E_n^λ if we truncate the series in (4.10) after $k = N - 1$. Proposition 4.1 shows that it is convenient to take $N > n + \lambda$. We consider the case $\lambda > 0$ (for negative λ, see Remark at the end of this Section):

PROPOSITION 4.3. *Let $n \in \mathbb{N}$, $\lambda > 0$, $\lambda \notin \mathbb{N}$, and $N \in \mathbb{N}$, $N > n + \lambda$. Then*

$$|R_n^\lambda(N)| \leq \mathcal{F}_n^\lambda(N) := \frac{n(n+\lambda)}{N} \sum_{j=0}^n \frac{1}{(n-j)!\,j!} \frac{(2\lambda+j)_n}{j+\lambda} \frac{|(-j-\lambda+1)_{N-1}|}{(j+\lambda+1)_{N-1}}. \tag{4.15}$$

Proof. Let us denote

$$A_{j,n}^\lambda = (-1)^j \binom{n}{j} \frac{n+\lambda}{n!} \frac{(2\lambda+j)_n}{j+\lambda} = (-1)^j \frac{n+\lambda}{j+\lambda} \frac{(2\lambda+j)_n}{(n-j)!\,j!}. \tag{4.16}$$

Then by (4.12),

$$R_n^\lambda(N) = 2n \sum_{k=N}^\infty m_{2k,n}^\lambda \frac{c_{2k,n}^\lambda}{k} = 2n \sum_{j=0}^n A_{j,n}^\lambda \sum_{k=N}^\infty \frac{(-j-\lambda)_k}{(j+\lambda+1)_k} \frac{c_{2k,n}^\lambda}{k},$$

so that

$$|R_n^\lambda(N)| \leq 2n \sum_{j=0}^n |A_{j,n}^\lambda| \sum_{k=N}^\infty \left| \frac{(-j-\lambda)_k}{(j+\lambda+1)_k} \right| \left| \frac{c_{2k,n}^\lambda}{k} \right| \leq \frac{2n}{N} \sum_{j=0}^n |A_{j,n}^\lambda| \sum_{k=N}^\infty \left| \frac{(-j-\lambda)_k}{(j+\lambda+1)_k} \right|. \tag{4.17}$$

Taking into account that

$$|A_{j,n}^\lambda| = (-1)^j A_{j,n}^\lambda, \quad \text{and} \quad |(-j-\lambda)_k| = (-1)^{[j+\lambda-1]} (-j-\lambda)_k, \text{ for } k \geq n + [\lambda] + 1,$$

where $[\lambda]$ stands for the integer value of λ, we get from (4.17),

$$|R_n^\lambda(N)| \leq \frac{2n(-1)^{[\lambda+1]}}{N} \sum_{j=0}^n A_{j,n}^\lambda \sum_{k=0}^\infty \frac{(-j-\lambda)_{k+N}}{(j+\lambda+1)_{k+N}}.$$

But

$$\sum_{k=0}^M \frac{(-j-\lambda)_{k+N}}{(j+\lambda+1)_{k+N}} = -\frac{1}{2} \frac{(-j-\lambda+1)_{N-1}}{(j+\lambda+1)_{N-1}} + \frac{(-j-\lambda+1)_{N+M}}{2\,(j+\lambda+1)_{N+M}},$$

so that

$$|R_n^\lambda(N)| \leq \frac{n(-1)^{[\lambda]}}{N} \sum_{j=0}^n A_{j,n}^\lambda \frac{(-j-\lambda+1)_{N-1}}{(j+\lambda+1)_{N-1}} = \mathcal{F}_n^\lambda(N), \tag{4.18}$$

and the statement follows. $\square$

Given $\varepsilon > 0$ we can use (4.15) in order to find a (preferably, lowest) value $N_0 \in \mathbb{N}$ such that $|R_n^\lambda(N)| \leq \varepsilon$. Obviously, the lower bound for N_0 will be $N_0 = n + [\lambda] + 1$. It is helpful to get also an upper bound for such an N_0, that can be obtained taking advantage of the geometric decay of $\mathcal{F}_n^\lambda(N)$. It is based on the following

PROPOSITION 4.4. *Let $n \in \mathbb{N}$, $\lambda > 0$, $\lambda \notin \mathbb{N}$, and $N \in \mathbb{N}$, $N > n + \lambda$. Then for all $h \in \mathbb{N}_0$,*

$$\mathcal{F}_n^\lambda(N+h) \leq \frac{F_n^\lambda(N)}{(N+\lambda+h)^{2\lambda}}, \tag{4.19}$$

where $\mathcal{F}_n^\lambda(N)$ is defined in (4.15), and

$$F_n^\lambda(N) = |\lambda \sin(\pi\lambda)|\,\Gamma^2(\lambda)\left(\frac{N-\lambda}{N+\lambda}\right)^{N-\lambda-1}\frac{n\,e^{2\lambda}}{\pi N}\sum_{j=0}^n \frac{n+\lambda}{j+\lambda}\frac{(2\lambda+j)_n}{(n-j)!\,j!}\frac{(\lambda)_j(\lambda+1)_j}{(N-\lambda-j)_j\,(N+\lambda)_j} \tag{4.20}$$

is a decreasing function in N, such that

$$F_n^\lambda(N) = O\left(\frac{1}{N}\right),\quad N\to\infty.$$

Proof. We can use the identity

$$\frac{(-j-\lambda+1)_{N+h-1}}{(j+\lambda+1)_{N+h-1}} = \frac{(-j-\lambda+1)_j\,(\lambda+1)_j}{(N+h-\lambda-j)_j\,(N+h+\lambda)_j}\frac{(-\lambda+1)_{N+h-1}}{(\lambda+1)_{N+h-1}}$$
$$= (-1)^j\,\frac{(\lambda)_j(\lambda+1)_j}{(N+h-\lambda-j)_j\,(N+h+\lambda)_j}\frac{(-\lambda+1)_{N-1}}{(\lambda+1)_{N-1}}\frac{(N-\lambda)_h}{(N+\lambda)_h}\,;$$

thus, with notation (4.16) and by (4.18),

$$|R_n^\lambda(N+h)| \le \frac{n(-1)^{[\lambda]}}{N}\frac{(-\lambda+1)_{N-1}}{(\lambda+1)_{N-1}}\frac{(N-\lambda)_h}{(N+\lambda)_h}\sum_{j=0}^n(-1)^j\,A_{j,n}^\lambda\frac{(\lambda)_j(\lambda+1)_j}{(N+h-\lambda-j)_j\,(N+h+\lambda)_j}$$
$$\le \frac{n(-1)^{[\lambda]}}{N}\frac{(-\lambda+1)_{N-1}}{(\lambda+1)_{N-1}}\frac{(N-\lambda)_h}{(N+\lambda)_h}\sum_{j=0}^n(-1)^j\,A_{j,n}^\lambda\frac{(\lambda)_j(\lambda+1)_j}{(N-\lambda-j)_j\,(N+\lambda)_j}$$
$$= \frac{(N-\lambda)_h}{(N+\lambda)_h}B_n^\lambda(N), \tag{4.21}$$

with

$$B_n^\lambda(N) = \frac{n}{N}\frac{|(-\lambda+1)_{N-1}|}{(\lambda+1)_{N-1}}\sum_{j=0}^n\frac{n+\lambda}{j+\lambda}\frac{(2\lambda+j)_n}{(n-j)!\,j!}\frac{(\lambda)_j(\lambda+1)_j}{(N-\lambda-j)_j\,(N+\lambda)_j}$$
$$= \frac{n}{N}\frac{\Gamma(N-\lambda)\Gamma^2(\lambda)|\lambda\sin(\pi\lambda)|}{\pi\Gamma(N+\lambda)}\sum_{j=0}^n\frac{n+\lambda}{j+\lambda}\frac{(2\lambda+j)_n}{(n-j)!\,j!}\frac{(\lambda)_j(\lambda+1)_j}{(N-\lambda-j)_j\,(N+\lambda)_j}, \tag{4.22}$$

where we have used again the identity $\Gamma(\lambda)\Gamma(1-\lambda)\sin(\pi\lambda) = \pi$. Alternatively, $B_n^\lambda(N)$ can be represented in terms of the following truncating and balanced hypergeometric series, valid for $\lambda \notin \mathbb{Z}$,

$$B_n^\lambda(N) = \frac{1}{N}\frac{|(-\lambda+1)_{N-1}|}{(\lambda+1)_{N-1}}\frac{n+\lambda}{(n-1)!}\frac{(2\lambda)_n}{\lambda}\,{}_4F_3\left(\begin{array}{c}-n,\,n+2\lambda,\,\lambda,\,\lambda\\2\lambda,-N+\lambda+1,N+\lambda\end{array}\middle|\,1\right).$$

In order to simplify the expression of the error, we may use (cf. [24, p.17]) that for $x \ge y \ge 1$,

$$\frac{\Gamma(y)}{\Gamma(x)} \le \frac{y^{y-1}\,e^x}{x^{x-1}\,e^y}\,.$$

Hence,

$$\frac{(N-\lambda)_h}{(N+\lambda)_h} = \frac{\Gamma(N+\lambda)}{\Gamma(N-\lambda)}\frac{\Gamma(N-\lambda+h)}{\Gamma(N+\lambda+h)} \le \frac{\Gamma(N+\lambda)}{\Gamma(N-\lambda)}\left(\frac{N-\lambda+h}{N+\lambda+h}\right)^{N-\lambda+h-1}\frac{e^{2\lambda}}{(N+\lambda+h)^{2\lambda}}\,.$$

It is straightforward to verify that $(x/(x+t))^{x-1}$ is decreasing in x for $x, t > 0$, so that

$$\frac{(N-\lambda)_h}{(N+\lambda)_h} \leq \frac{\Gamma(N+\lambda)}{\Gamma(N-\lambda)} \left(\frac{N-\lambda}{N+\lambda}\right)^{N-\lambda-1} \frac{e^{2\lambda}}{(N+\lambda+h)^{2\lambda}} .$$

Gathering this inequality and (4.22) in (4.21) we obtain the statement of the Proposition. $\square$

COROLLARY 4.5. *Let $\varepsilon > 0$, $n \in \mathbb{N}$, $\lambda > 0$, $\lambda \notin \mathbb{N}$, and $N \in \mathbb{N}$, $N > n + \lambda$. Then if for $F_n^\lambda(N)$ defined in (4.20),*

$$h \geq \max \left\{ \left[\left(\frac{F_n^\lambda(N)}{\varepsilon}\right)^{\frac{1}{2\lambda}} - N - \lambda\right], 0 \right\}, \qquad (4.23)$$

then

$$\left|\mathcal{F}_n^\lambda(N+h)\right| \leq \varepsilon .$$

Hence, if we want to find a suitable value of $N_0 \in \mathbb{N}$ for which $\left|R_n^\lambda(N_0)\right| \leq \varepsilon$, we can use the following procedure:

ALGORITHM 3

(1) Take $N = n + [\lambda] + 1$ and compute $\mathcal{F}_n^\lambda(N)$;

(2) if $\mathcal{F}_n^\lambda(N) \leq \varepsilon$, put $N_0 = N$ and quit;

(3) else compute h equal to the r.h.s. of (4.23);

(4) if $h > 0$, use bisection in order to find the lowest $N_0 \in [N, N+h] \cap \mathbb{N}$ such that $\mathcal{F}_n^\lambda(N_0) \leq \varepsilon$

Obviously, we can use a more sophisticated zero-finding method in the procedure above; however, usually bisection, which is simple and easy to implement, is sufficient for our needs.

REMARK: A simple alternative for truncation of the series in (4.10) can be computing E_n^λ for different integer values of λ (say, $[\lambda]$ and $[\lambda]+1$) and interpolating the value of E_n^λ for the non-integer λ. Nevertheless, as numerical experiments in Section 5 show, this approach is not very satisfactory. For small values of λ it yields large errors, and for large λ's the loss in speed computing the entropy at least twice can be compensated by larger truncation values N.

REMARK: For small values of the parameter λ the error bounds above usually yield large truncation values N. This occasionally might justify the use of explicit formulas (4.12) and (4.13) for computation of $m_{2k,n}^\lambda$ instead of Step 3. As λ grows larger, the explicit evaluation of Pochhammer's symbols rapidly becomes substantially more time consuming and less accurate than matrix multiplication.

Furthermore, for $-1/2 < \lambda < 0$ the structure of the coefficients $m_{2k,n}^\lambda$ yields extremely pessimistic upper bounds for the error $|R_n^\lambda(N)|$, with a rate of convergence even lower than $1/N$ established in Theorem 2.1. Nevertheless, in this case the Gegenbauer polynomials are uniformly bounded on $[-1, 1]$, and truncation error is estimated better using formulas (2.6)–(2.7).

5. Numerical experiments for Gegenbauer polynomials. In this section we discuss briefly the performance of the algorithm presented above, and compare it with some alternative algorithms used for computing the entropy of Gegenbauer polynomials.

We will compute E_n^λ for several values of λ. In all cases Algorithm 2 was implemented in Matlab[TM] and executed on a computer with a single AMD Athlon[TM] XP2000+ processor, 256 Mb RAM, and running Matlab 6 under Windows. In particular, specific routines for sparse matrix construction have been used through
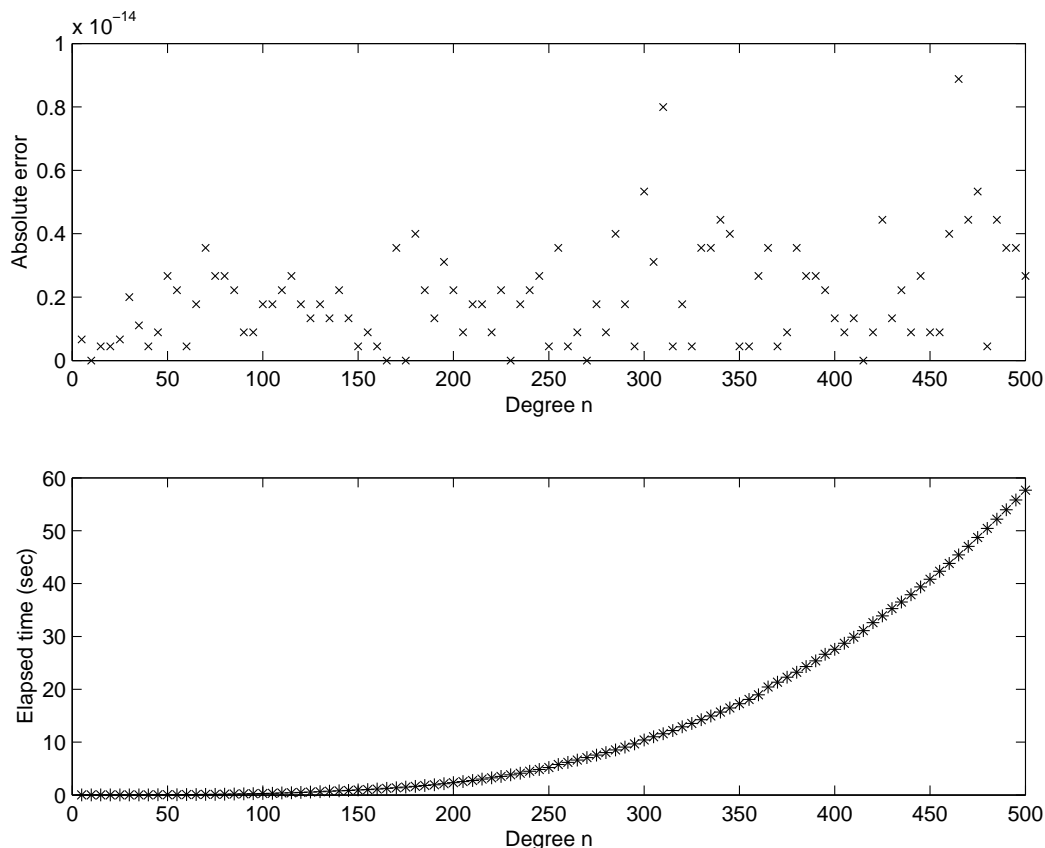
FIG. 5.1. *Error and execution time of the algorithm for computation of E_n^2.*

Matlab built in functions `spdiags` and `speye`. For the experiments no special performance optimization techniques have been used, although we implemented vectorization when available.

One obvious test situation corresponds to $\lambda = 0, 1, 2$, when the explicit value of the entropy is known (cf. formulas (4.5)–(4.6)). Figure 5.1 shows that the error of the algorithm (comparing with the exact value (4.6)) is negligible. The execution time, computed as an average of 100 runs of the algorithm, grows geometrically with the degree n (due to the proportional growth of the matrices involved), but still it is below 1 minute for n as large as 500.

It is also illustrating to compare E_n^λ with their asymptotic expansion (4.7) truncated after the first and the second terms (Fig. 5.2).

In order to illustrate the performance of Algorithm 2 we compare it with the following procedures to compute E_n^λ:
- By formula (4.4) using adaptative quadrature implemented in Mathematica™ 4.2;
- By formula (4.4) using functions `quad` and `quadl` of Matlab 6, applied to explicit expressions of the polynomials G_n^λ with coefficients computed using Mathematica 4.2 with exact arithmetics;
- By the algorithm described in [9]. This approach is valid for integer values of λ only.

In Table 5.1 we compare the errors and execution times of the procedures above with those of Algorithm 2. The execution time is taken as the average of 100 runs of the corresponding algorithms.

As it was mentioned above, the bound in (4.15) usually overestimates the error. For that purpose we truncate first the series in (4.10) at N such that $\mathcal{F}_n^\lambda(N)$ is not greater than the machine epsilon; the corre-
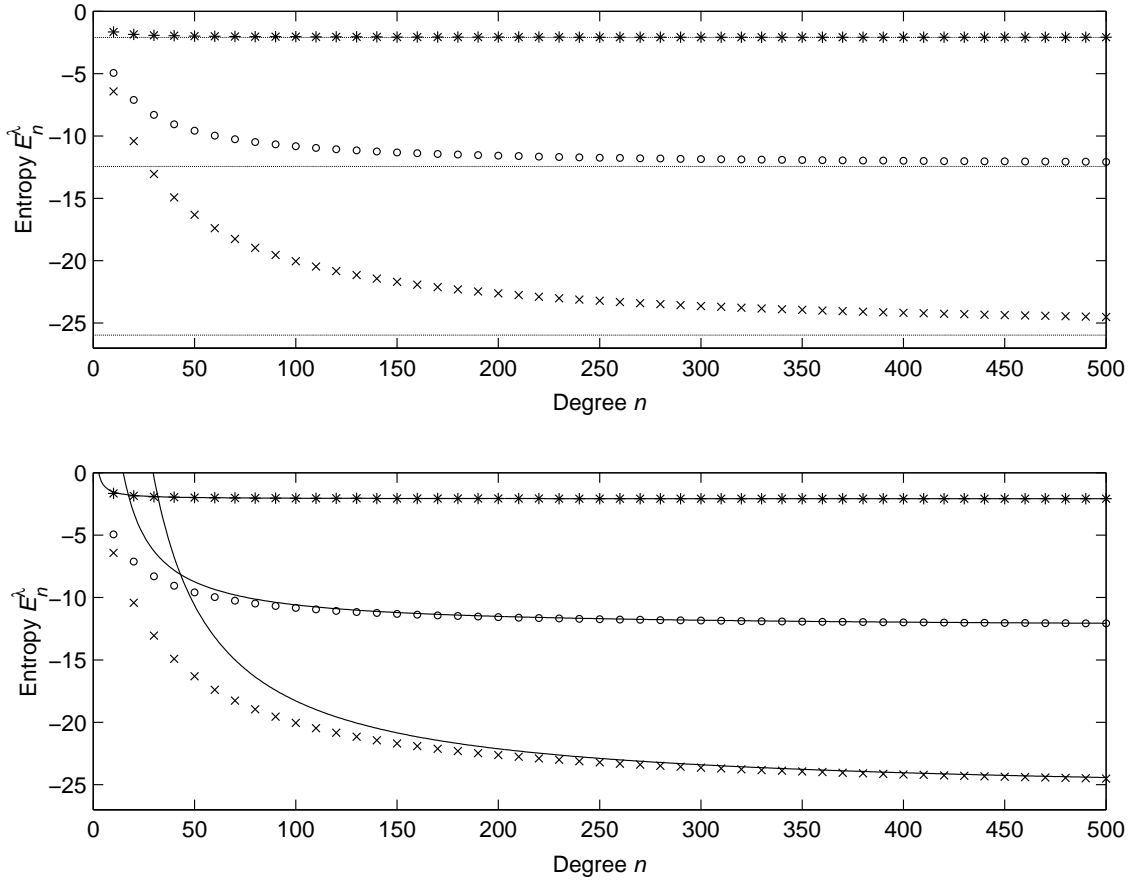
FIG. 5.2. *Entropy E_n^λ for $\lambda = 2$ (marked with '*'), $\lambda = 10$ (marked with 'o') and $\lambda = 20$ (marked with 'x') compared with its limit $\mathcal{E}_0^\lambda$ (top) and the asymptotic expression $\mathcal{E}_0^\lambda + \mathcal{E}_1^\lambda / n$ (bottom).*

sponding value of E_n^λ is assumed as the "true" value of the entropy. We compare it with the approximation of E_n^λ that we obtain if we truncate the series in (4.10) at N given by Algorithm 3 (Fig. 5.3).

Finally, it is tempting to avoid the question of truncation error in (4.10) by applying Algorithm 2 to $\lambda \in \mathbb{N}_0$ and computing E_n^λ for non-integer values of the parameter by interpolation. As experiment, we compute E_{200}^λ for half-integer values of λ using two strategies. In Figure 5.4, we observe the execution time (in seconds) for truncating the series in (4.10) in the way that $\mathcal{F}_{200}^\lambda \leq 10^{-6}$ (dots), and for interpolating E_{200}^λ by cubic splines using the values of the entropy for $\lambda \pm 1/2$, $\lambda \pm 3/2$ (diamonds). Asterisks represent the errors of interpolation. As we see, interpolation usually yields large errors for small values of λ, where it still could be competitive, since for large λ's we are penalized by the time invested in computing the entropy at least twice.

6. Computation of the entropy of spherical harmonics. In this section, to show the usefulness of our computational algorithm as well as the close connection of the entropy of Gegenbauer polynomials analyzed in detail in the three previous sections, we determine the spatial complexity of some quantum-mechanical prototype and real systems with central potentials (rigid rotator, harmonic oscillator, hydrogen atom, Rydberg atoms, some diatomic molecules, etc.) by means of the entropy of the spherical harmonics,

$$S_{l,m}[Y] := -\int |Y_{l,m}(\Omega)|^2 \ln |Y_{l,m}(\Omega)|^2 \, d\Omega,$$

Method	Absolute error				Time (sec)			
	$n = 10$	$n = 25$	$n = 50$	$n = 100$	$n = 10$	$n = 25$	$n = 50$	$n = 100$
(i)	4.0×10^{-7}	1.4×10^{-4}	5.8×10^{-4}	1.4×10^{-3}	0.30	0.55	1.13	2.51
(ii)	4.3×10^{-6}	1.3×10^{-5}	5.7×10^{-5}	1.1×10^{-2}	0.08	0.28	0.89	2.61
(iii)	7.7×10^{-8}	1.4×10^{-7}	8.6×10^{-7}	5.2×10^{-6}	0.11	0.31	0.74	3.01
(iv)	0	0	1.8×10^{-26}	1.3×10^{-7}	0.006	0.008	0.01	0.02
(v)	4.4×10^{-15}	2.2×10^{-16}	6.0×10^{-15}	2.7×10^{-15}	0.002	0.007	0.03	0.25

TABLE 5.1

Absolute error (left half) and time of computation (in seconds) of E_n^2 for $n = 10, 25, 50, 100$ by means of (i) adaptative quadrature of Mathematica 4.1 with extended precision (left half) or using floating point arithmetics (right half), (ii) `quad` *and (iii)* `quadl` *functions of Matlab, (iv) algorithm from [9] implemented in Mathematica 4.1 with extended precision, and (v) Algorithm 2.*

where $\Omega \equiv (\theta, \varphi)$, $d\Omega = \sin\theta \, d\theta \, d\varphi$, with $0 \leq \theta \leq \pi$ and $0 \leq \varphi \leq 2\pi$, and $Y_{l,m}(\Omega)$ denotes the spherical harmonics which depend on the orbital and azimuthal quantum numbers, l and m respectively. It is well-known that the principal quantum number $n \in \mathbb{N}_0$, together with $l \in \mathbb{N}_0$ and $m \in \mathbb{Z}$, completely characterize a single-particle system with a central potential. Moreover, for a given n the orbital quantum numbers $l \leq n - 1$, and for a given l the azimuthal quantum number $-l \leq m \leq +l$.

The spatial or angular wavefunction of the system, which defines its bulky shape, can be expressed in terms of Gegenbauer polynomials as (cf. [36])

$$Y_{lm}(\Omega) = N_{l,m} e^{im\varphi} (\sin\theta)^{|m|} C_{l-|m|}^{|m|+\frac{1}{2}} (\cos\theta),$$

where C_n^λ are Gegenbauer polynomials normalized as in (4.2), and

$$N_{l,m} = \left[\frac{\left(l + \frac{1}{2}\right)(l - |m|)! \left[\Gamma\left(|m| + \frac{1}{2}\right)\right]^2}{2^{1-2|m|}\pi^2 (l + |m|)!} \right]^{1/2}.$$

Then, taking into account relation (4.3), the entropy $S_{l,m}[Y]$ is expressed in terms of the entropy of the Gegenbauer orthonormal polynomials, defined in (4.4), as

$$S_{l,m}[Y] = \ln\left(\frac{2\pi}{c_{|m|+1/2}}\right) + E_{l-|m|}^{|m|+1/2} - |m| \left[2\psi(l + |m| + 1) - 2\psi\left(l + \frac{1}{2}\right) - 2\ln 2 - \frac{1}{l+1/2}\right],$$

where c_λ is defined in (4.1).

Thus, we can apply Algorithm 2 in order to compute the entropy of the spherical harmonics $S_{l,m}[Y]$ for different values of the quantum numbers l and m. In Fig. 6.1, values of $S_{200,m}[Y]$ are computed for integer values of $0 \leq m \leq 200$. This figure illustrates that for a given l the entropy is higher around the center of the manifold of azimuthal quantum numbers $m = -l, -l + 1, -l + 2, \ldots, l - 1, l$, than at its extremes, indicating that the spherical harmonics are much more localized for the largest values of $|m|$. Moreover, the entropy is approximately constant in the interval $-l/2 \lesssim m \lesssim l/2$, and then it monotonically decreases when $|m|$ grows up to its largest allowed value l. The origin of this intriguing phenomenon is the delicate interplay of the sinus factor and the Gegenbauer polynomial involved in the spherical harmonics, which deserves further numerical investigation.
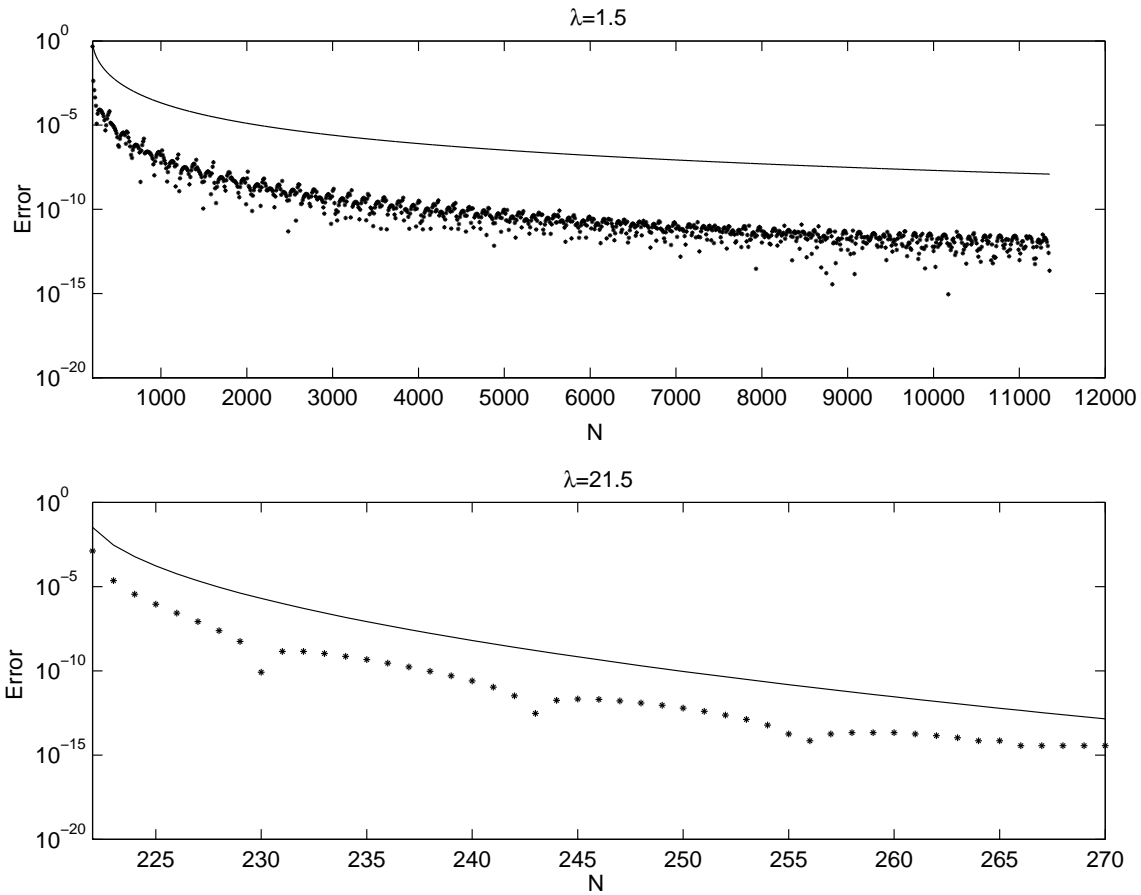
FIG. 5.3. *Error in computing* E_{200}^{λ} *by (4.10) truncating the series at N (dots) compared with the error bound* $\mathcal{F}_{200}^{\lambda}(N)$ *(solid line) for $\lambda = 1.5$ (top) and $\lambda = 21.5$ (bottom).*

REFERENCES

[1] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of Mathematical Functions*, Dover Publ., New York, 1972.
[2] G. E. ANDREWS, R. ASKEY, AND R. ROY, *Special Functions*, vol. 71 of Encyclopedia of Math. and its Applic., Cambridge Univ. Press, April 1999. ISBN 0-521-62321-9.
[3] J. ANGULO, J. ANTOLIN, J. C. CUCHI, AND A. ZARZO, *Maximum entropy technique with logarithmic constraints: Estimation of atomic radial densities*, Eur. Phys. J. D, 7 (1999), pp. 479–485.
[4] 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 (1995), pp. 373–395.
[5] W. BECKNER, *Inequalities in Fourier analysis*, Ann. Math., 102 (1975), pp. 159–182.
[6] I. BIALYNICKI-BIRULA AND J. MYCIELSKY, *Uncertainty relations for information entropy in wave mechanics*, Commun. Math. Phys., 44 (1975), pp. 129–132.
[7] V. S. BUYAROV, *On information entropy of Gengenbauer polynomials*, Vesnt. Mosk. Univ., Ser. 1, 6 (1997), pp. 8–11. In Russian.
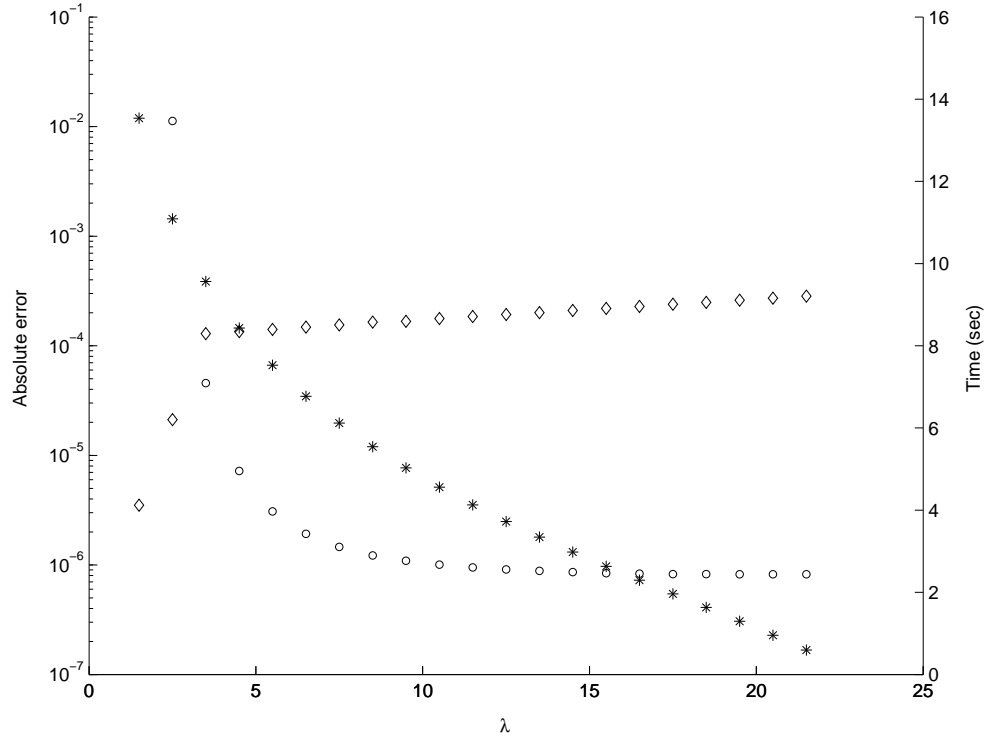
FIG. 5.4. *Comparison of the interpolation strategy vs. truncation of the series in (4.10).*

[8] 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 (1999), pp. 153–166.

[9] V. S. BUYAROV, P. LÓPEZ-ARTÉS, A. MARTÍNEZ-FINKELSHTEIN, AND W. VAN ASSCHE, *Information entropy of Gegenbauer polynomials*, J. Phys. A, 33 (2000), pp. 6549–6560.

[10] D. C. CHAMPENEY, *A Handbook of Fourier Theorems*, Cambridge Univ. Press, Cambridge, UK, 1987.

[11] J. S. DEHESA, A. MARTÍNEZ-FINKELSHTEIN, AND J. SÁNCHEZ-RUIZ, *Quantum information entropies and orthogonal polynomials*, J. Comput. Appl. Math., 133 (2001), pp. 23–46.

[12] J. S. DEHESA, W. VAN ASSCHE, AND R. J. YÁÑEZ, *Information entropy of classical orthogonal polynomials and their application to the harmonic oscillator and Coulomb potentials*, Methods and Appl. Analysis, 4 (1997), pp. 91–110.

[13] J. S. DEHESA, R. J. YAÑEZ, A. I. APTEKAREV, AND V. S. BUYAROV, *Strong asymptotics of Laguerre polynomials and information entropies of 2D harmonic oscillator and 1D Coulomb potentials*, J. Math. Physics, 39 (1998), pp. 3050–3060.

[14] A. ERDÉLYI, ed., *Higher Transcendental Functions*, vol. I, McGraw-Hill, New York, 1953.

[15] S. R. GADRE, *Information theoretical approaches to quantum chemistry*, in Reviews of Modern Quantum Chemistry, K. D. Sen, ed., Dordrecht, 2002, Kluwer.

[16] S. R. GADRE AND R. D. BENDALE, *Maximization of atomic information entropy sum in configuration and momentum spaces*, Int. J. Quantum Chem., 28 (1985), pp. 311–319.

[17] H. GRAD, *The many faces of entropy*, Comm. Pure Appl. Math., 14 (1961), pp. 323–354.

[18] I. S. GRADSHTEIN AND I. M. RYZHIK, *Table of Integrals, Series and Products*, Academic Press, San Diego, CA, fifth ed., 1995.

[19] W. T. GRANDY, *Resource letter ITP-1: Information theory in physics*, Amer. J. Phys., 65 (1997), pp. 466–476.

[20] A. S. HOLEVO, *Quantum coding theorems*, Russian Math. Surveys, 53 (1998), pp. 1295–1331.

[21] E. T. JAYNES, *Information theory and statistical mechanics*, Phys. Rev., 106 (1957), pp. 620–630.

[22] A. JEFFREY, *Handbook of Mathematical Formulas and Integrals*, Academic Press, San Diego, CA, 1995.

[23] S. KARLIN AND Y. RINOTT, *Entropy inequalities for classes of probability distributions I. the univariate case*, Adv. Appl. Prob., 13 (1981), pp. 93–112.

[24] Y. L. LUKE, *Mathematical Functions and their Approximations*, Academic Press, New York, 1975.

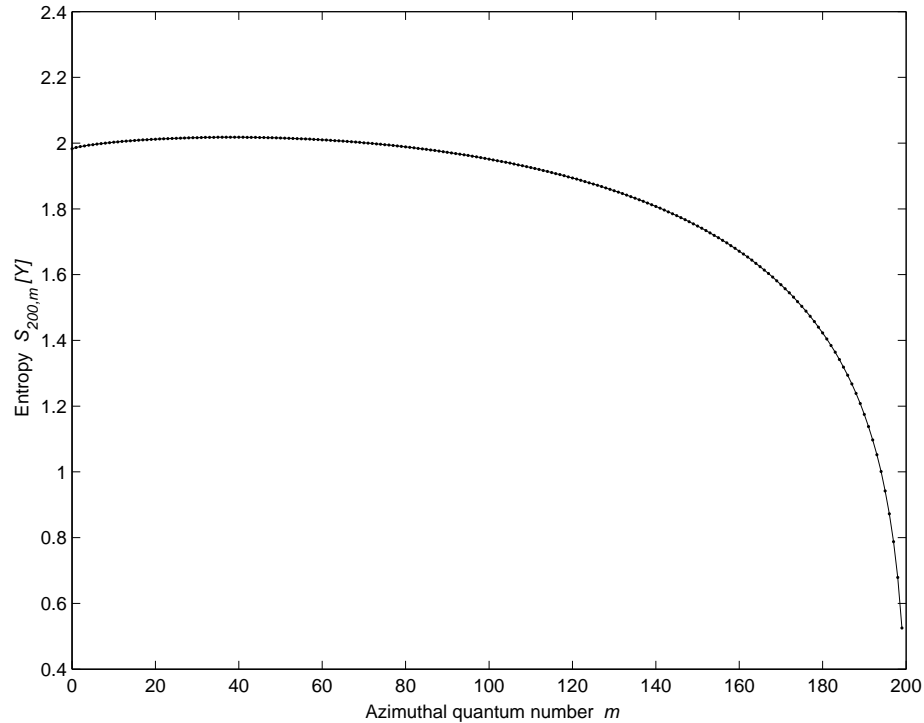[25] M. A. NIELSEN AND I. L. CHUANG, *Quantum Computation and Quantum Information*, Cambridge University Press,

Fig. 6.1. *Entropy $S_{200,m}[Y]$ for azimuthal quantum numbers $m = 0, 1, \ldots, 199$.*

Cambridge, 2000.

[26] M. OHYA AND D. PETZ, *Quantum Entropy and its Use*, Springer Verlag, New York, 1993.

[27] R. G. PARR AND W. YANG, *Density Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.

[28] E. A. RAKHMANOV, *On the asymptotics of the ratio of orthogonal polynomials*, Math. USSR Sb., 32 (1977), pp. 199–213.

[29] J. F. SÁNCHEZ LARA, *On the asymptotic expansion of the entropy of Gegenbauer polynomials*, J. Comput. Appl. Math., 142 (2002), pp. 401–409.

[30] C. E. SHANNON, *A mathematical theory of communication*, Bell Syst. Tech. J., 27 (1948), pp. 379–423, 623–656. Reprinted in C.E. Shannon, Claude Elwood Shannon: Collected Papers, IEEE Press, New York, 1993.

[31] ———, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana-Campaign, 1949.

[32] G. SZEGŐ, *Orthogonal Polynomials*, vol. 23 of Amer. Math. Soc. Colloq. Publ., Amer. Math. Soc., Providence, RI, fourth ed., 1975.

[33] L. SZILARD, *On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings*, Z. Phys., 53 (1929), pp. 840–856. Translation by A. Rapoport and M. Knoller, reprinted in Quantum Theory and Measurement, J.A. Wheeler and W.H. Zurek, eds., Princeton Univ. Press, Princeton, 1983.

[34] W. VAN ASSCHE, *Asymptotics for orthogonal polynomials*, vol. 1265 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1987.

[35] R. J. YÁÑEZ, W. VAN ASSCHE, AND J. S. DEHESA, *Position and momentum information entropies of the D-dimensional harmonic oscillator and hydrogen atom*, Physical Rev. A, 50 (1994), pp. 3065–3079.

[36] R. J. YAÑEZ, W. VAN ASSCHE, R. GONZÁLEZ-FÉREZ, AND J. S. DEHESA, *Entropic integrals of hyperspherical harmonics and spatial entropy of D-dimensional central potentials*, J. Math. Physics, 40 (1999), pp. 5675–5686.

[37] A. ZARZO, J. ANGULO, J. C. CUCHI, AND J. ANTOLIN, *Maximum entropy and Padé-like approximations to atomic scattering factors*, Z. Physik D, 41 (1997), pp. 175–179.

[38] W. H. ZUREK, ed., *Complexity, Entropy and the Physics of Information*, Redwood, 1990, Santa Fe Institute Studies in the Sciences of Complexity, Addison-Wesley.