

A study of the compactness of wave functions based on Shannon entropy indices: a seniority number approach

Luis Lain¹ · Alicia Torre¹ · Diego R. Alcoba^{2,3} · Ofelia B. Oña⁴ · Gustavo E. Massaccesi⁵

Received: 16 April 2015 / Accepted: 16 June 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract This work reports the formulation of Shannon entropy indices in terms of seniority numbers of the Slater determinants expanding an N -electron wave function. Numerical determinations of those indices prove that they provide a suitable quantitative procedure to evaluate compactness of wave functions and to describe their configurational structures. An analysis of the results, calculated for full configuration interaction wave functions in selected atomic and molecular systems, allows one to compare and to discuss the behavior of several types of molecular orbital basis sets in order to achieve more compact wave

function expansions, and to study their multiconfigurational character.

Keywords Compactness of wave functions · Seniority number · Shannon entropy indices · Optimization of molecular orbital basis sets

1 Introduction

The configuration interaction (CI) methods have played an important role in describing N -electron systems, since they demand a lower computational cost than that required for determining the full configuration interaction (FCI) expansions which provide the exact descriptions for a given Hilbert space. Consequently, there has been a considerable interest in formulating N -electron wave functions in terms of CI expansions providing a rapid convergence to the FCI ones [1–4]. Traditionally, the CI wave functions have been expanded by means of N -electron Slater determinants selected according to their excitation degrees with respect to a given reference determinant. However, more recently, another selection criterion has also been proposed. This criterion is based on the seniority number of the Slater determinants used to construct the CI expansions [5–11]. Results arising from both excitation- and seniority number-based CI schemes show that the seniority number-based selection procedure is particularly suitable to describe systems exhibiting strong (static) correlation [5], what has increased the interest on this approach [9–11]. As is well known, the seniority number of a Slater determinant is defined as the number of singly occupied orbitals which possesses that determinant [12, 13]. The seniority number concept has been extended in Refs. [6, 9] to N -electron wave functions which describe electronic states of atomic and molecular systems,

Dedicated to Prof. P. R. Surjan on occasion of his 60th birthday.

Published as part of the special collection of articles “Festschrift in honour of P. R. Surjan.”

✉ Luis Lain
qfplapel@lg.ehu.es

- ¹ Departamento de Química Física, Facultad de Ciencia y Tecnología, Universidad del País Vasco, Apdo. 644, 48080 Bilbao, Spain
- ² Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, 1428 Buenos Aires, Argentina
- ³ Instituto de Física de Buenos Aires, Consejo Nacional de Investigaciones Científicas y Técnicas, Ciudad Universitaria, 1428 Buenos Aires, Argentina
- ⁴ Instituto de Investigaciones Físicoquímicas Teóricas y Aplicadas, Universidad Nacional de la Plata, CCT La Plata, Consejo Nacional de Investigaciones Científicas y Técnicas, Diag. 113 y 64 (s/n), Sucursal 4, CC 16, 1900 La Plata, Argentina
- ⁵ Departamento de Ciencias Exactas, Ciclo Básico Común, Universidad de Buenos Aires, Ciudad Universitaria, 1428 Buenos Aires, Argentina

as well as to N -electron spin-adapted Hilbert spaces. The expectation value of the seniority number operator with respect to an N -electron wave function is a weighted sum of the seniority numbers of all determinants involved in the expansion of that wave function. The weights that determine those expectation values depend on the molecular orbital basis set used to express the wave function, and consequently, this feature has been utilized to evaluate the compactness of the FCI and CI expansions in several molecular basis sets. Likewise, the seniority number value with respect to a wave function allows one to analyze the multiconfigurational character of the N -electron expansion, which is useful to describe the static and dynamic correlation of a determined state [14–17].

On the other hand, in Ref. [18], the extent of the multiconfigurational character of an N -electron wave function was evaluated by means of numerical determinations of an index set formulated within the Shannon information entropy approach [19–21]. This treatment provides a suitable information concerning the distribution of the wave function among different configurations characterized by the excitation degree of the Slater determinants. The aim of this work is to extend this methodology to the seniority-based CI scheme and to report the corresponding Shannon index numerical values in terms of the contributions of Slater determinants classified according to the seniority number criterion. More recently, in Ref. [6], we have proposed unitary transformations which lead to the construction of basis sets of molecular orbitals in which the expectation values of the seniority number operator with respect to N -electron wave functions reach minimum values. The results found using this type of molecular orbitals show that the wave function expansions present a more rapid convergence than those arising from the use of other molecular orbitals [9, 11]. Another aim of this work is to evaluate and compare quantitatively, by means of the proposed Shannon entropy indices, the compactness of wave functions expressed in canonical molecular orbital (CMO) basis sets, natural orbitals (NO), and those mentioned orbitals M_{\min} , which minimize the expectation value of the seniority number operator.

This article has been organized as follows. Section 2 summarizes the notation and formulation of the main concepts used in this work; it also reports the formulation of the Shannon entropy indices in terms of the seniority numbers of the Slater determinants. In Sect. 3, we present numerical values of those indices for wave functions of selected atomic and molecular systems; these values allow one to characterize the compactness of the wave function expansions. The calculation level and the computational details are also indicated in this section. An analysis and discussion of these results are reported in Sect. 4. Finally, in the last section, we highlight the main conclusions and perspectives of this work.

2 Theoretical framework

The K orbitals of an orthonormal basis set will be denoted by i, j, k, l, \dots and their corresponding spin-orbitals by $i^\sigma, j^{\sigma'}, \dots$ (σ and σ' mean the spin coordinates α or β). The spin-free version of the N -electron seniority number operator $\hat{\Omega}$ has been formulated as [6, 9, 11]

$$\hat{\Omega} = \sum_{i=1}^K (\hat{E}_i^i - \hat{E}_{ii}^{ii}) \quad (1)$$

where $\hat{E}_i^i = \sum_{\sigma} a_{i\sigma}^\dagger a_{i\sigma}$ and $\hat{E}_{ii}^{ii} = \sum_{\sigma, \sigma'} a_{i\sigma}^\dagger a_{i\sigma'}^\dagger a_{i\sigma'} a_{i\sigma}$ are the spin-free first- and second-order replacement operators, respectively [22–25] and $a_{i\sigma}^\dagger$ and $a_{i\sigma}$ are the usual creation and annihilation fermion operators [26].

Closing both sides of Eq. (1) by an N -electron Slater determinant Λ of S_z spin projection quantum number, one obtains

$$\Omega = \langle \Lambda | \hat{\Omega} | \Lambda \rangle \quad (2)$$

where, according to Eq. (1), the expectation value Ω is the difference $\sum_i \langle E_i^i \rangle - \sum_i \langle E_{ii}^{ii} \rangle$, which is the number of total electrons N minus the number of electrons corresponding to doubly occupied orbitals. The possible values for the Ω parameter are positive integers belonging to the sequence $\Omega = 2|S_z|, 2(|S_z| + 1), \dots, \Omega_{\max}$ (where $\Omega_{\max} = N$ if $K \geq N$ and $\Omega_{\max} = 2K - N$ if $K < N$); their meaning is the number of non-repeated orbitals in each determinant. That Ω parameter allows one to classify the Slater determinants Λ of S_z quantum number, according to the corresponding seniority level, and they will be denoted hereafter by $\Lambda(\Omega)$. Consequently, a FCI N -electron wave function with given spin quantum numbers S and S_z will be expressed by

$$|\Psi(N, S, S_z)\rangle = \sum_{\Omega=2S}^{\Omega_{\max}} \sum_{\Lambda(\Omega)} C_{\Lambda(\Omega)} |\Lambda(\Omega)\rangle \quad (3)$$

where $C_{\Lambda(\Omega)}$ stands for the coefficient corresponding to the Slater determinant $\Lambda(\Omega)$. Obviously, since there is no contribution of Slater determinants with $\Omega < 2S$ to spin-adapted N -electron wave functions ($\langle \Lambda(\Omega < 2S) | \Psi(N, S, S_z) \rangle = 0$), the lowest integer in the sum \sum_{Ω} is $2S$. If we truncate the series \sum_{Ω} in Eq. (3), we obtain CI(Ω) wave function expansions involving only Slater determinants belonging to the selected Ω levels.

According to Eqs. (1) and (3), the expectation value of the operator $\hat{\Omega}$ with respect to the FCI wave function $\Psi(N, S, S_z)$ is [6]

$$\langle \hat{\Omega} \rangle_{\Psi(N, S)} = N - \sum_{\Omega=2S}^{\Omega_{\max}} \sum_{\Lambda(\Omega)} |C_{\Lambda(\Omega)}|^2 \sum_i^K \langle \Lambda(\Omega) | \hat{E}_{ii}^{ii} | \Lambda(\Omega) \rangle \quad (4)$$

which is a spin-free quantity, independent of the S_z value, and consequently, we have dropped that quantum number. The coefficients $C_{\Lambda(\Omega)}$ and the $\langle \hat{\Omega} \rangle_{\Psi(N,S)}$ values are strongly dependent on the molecular orbital set utilized to formulate the Slater determinants $\Lambda(\Omega)$ in the expansion of the wave function expressed by Eq. (3). As mentioned in the Introduction, in Refs. [6, 9, 11], we have performed unitary transformations of the molecular orbitals, based on iterative procedures [27], which lead to the minimization of the $\langle \hat{\Omega} \rangle_{\Psi(N,S)}$ values; the resulting molecular orbital basis sets have been denominated M_{\min} . That minimization requires the search of molecular orbitals leading to high values for the coefficients $|C_{\Lambda(\Omega)}|$ corresponding to the determinants which possess greater doubly occupied orbital numbers, i.e., those providing higher values of the $\sum_i^K \langle \Lambda(\Omega) | \hat{E}_{ii}^i | \Lambda(\Omega) \rangle$ quantities. Our results [6, 9, 11] have proven that the expansions for ground-state wave functions of atomic and molecular systems expressed in the molecular orbital basis sets M_{\min} turn out to be more compact than those arising from the canonical molecular orbitals (CMO) or natural orbitals (NO).

Quantitative measures of the compactness of an N -electron wave function have been reported in Ref. [18] by means of the informational content (I_C (or Shannon entropy) within the traditional CI expansion method based on Slater determinants classified according to the excitation level with respect to a given reference determinant. Assuming that the N -electron wave function is normalized to unity ($\sum_{\Omega} \sum_{\Lambda(\Omega)} |C_{\Lambda(\Omega)}|^2 = 1$), the counterpart formulation of that index for the seniority-based CI approach is

$$I_C = - \sum_{\Omega} \sum_{\Lambda(\Omega)} |C_{\Lambda(\Omega)}|^2 \log_2 |C_{\Lambda(\Omega)}|^2, \quad C_{\Lambda(\Omega)} \neq 0 \quad (5)$$

in which the index Ω runs over all the values defining the chosen CI (Ω) expansion seniority levels. According to Eq. (5), the I_C index accounts for the wave function configurational distribution, having a minimum value in case of a single-determinant wave function.

The values of this I_C index quantify the multiconfigurational character of the CI wave function but do not report any detailed information on the contributions corresponding to different seniority subspaces. For CI (Ω) expansions involving several values of the Ω index, we can define a weight W_{Ω} which groups the contributions of all the Slater determinants with given seniority number Ω in that expansion

$$W_{\Omega} = \sum_{\Lambda(\Omega)} |C_{\Lambda(\Omega)}|^2 \quad (6)$$

These weights provide the definition of the cumulative index I_W , which in the seniority number approach is

$$I_W = - \sum_{\Omega} W_{\Omega} \log_2 W_{\Omega}, \quad W_{\Omega} \neq 0 \quad (7)$$

which evaluates that entropic quantity in terms of the weights corresponding to the seniority numbers Ω , providing a measure of the distribution of the wave function on different seniority subspaces.

One can also consider the distribution of each Ω subspace in terms of its corresponding Slater determinants and calculate its specific entropic index, which can be evaluated by means of the relationship

$$I_{\Omega} = - \sum_{\Lambda(\Omega)} \frac{|C_{\Lambda(\Omega)}|^2}{W_{\Omega}} \log_2 \frac{|C_{\Lambda(\Omega)}|^2}{W_{\Omega}}, \quad C_{\Lambda(\Omega)} \neq 0 \quad (8)$$

where the denominators W_{Ω} have been introduced for normalization requirements. Formula (8) accounts for the configuration distribution within a determined seniority number level.

As mentioned above, the multiconfigurational character of an N -electron wave function expanded in terms of Slater determinants allows one to distinguish between systems exhibiting static (strong) correlation (in which a suitable zeroth-order description requires several Slater determinants) and those possessing dynamic correlation (in which a single Slater determinant is a good zeroth-order wave function). In the next sections, we report numerical values of the I_C and I_W indices in selected atomic and molecular systems, in order to assess the ability of these devices to describe quantitatively both types of electronic correlation within the seniority number approach. Likewise, we present values of the I_{Ω} index which show the influence of the bond stretching on the configurational distribution within the seniority number subspaces. As these Shannon entropy indices and the seniority number quantity for a determined wave function are not invariant under a unitary single-particle transformation, it is possible to perform molecular basis set unitary transformations and to compare values of these entropic indices according to the different molecular basis sets utilized. In particular, we compare values of Shannon indices arising from the molecular basis sets M_{\min} (in which the seniority number achieves their minimum values) with those provided by the CMO and NO sets.

3 Results

We have determined expansions of wave functions of several atomic and molecular systems in their ground states, at FCI level. These wave functions have been expressed in the three mentioned molecular basis sets CMO, NO, and M_{\min} , in order to study their compactness in different molecular orbital basis sets. Our aim is to analyze the structure and compactness of those expansions by means of the entropic indices proposed in Eqs. (5), (7), and (8) according to the seniority numbers. We have mainly chosen the systems of

Table 1 Calculated values of the I_C and I_W quantities (Eqs. (5) and (7)) for the ground states of atomic and molecular systems described by FCI expansions expressed in the canonical molecular orbitals(CMO), in the orbitals which minimize the seniority number M_{\min} and in the natural orbitals (NO)

System	I_C			I_W		
	CMO	M_{\min}	NO	CMO	M_{\min}	NO
Be(STO-3G)	0.649	0.648	0.648	0.001	0.000	0.000
Be(cc-pVDZ)	0.737	0.602	0.602	0.181	0.001	0.001
Mg(6-31G)	0.648	0.521	0.522	0.184	0.007	0.007
LiH(R_e)	0.231	0.167	0.167	0.081	0.001	0.001
LiH(R_{st})	1.709	0.864	0.864	0.678	0.002	0.002
BeH ⁺ (R_e)	0.222	0.164	0.164	0.073	0.001	0.001
BeH ⁺ (R_{st})	1.767	0.990	0.990	0.733	0.001	0.001
Li ₂ (R_e)	0.664	0.598	0.598	0.078	0.003	0.003
Li ₂ (R_{st})	1.103	0.854	0.854	0.235	0.003	0.003
BH(R_e)	0.681	0.552	0.556	0.169	0.009	0.015
BH(R_{st})	0.979	0.802	1.005	0.224	0.026	0.028
BH ₂ ⁺ (R_e)	0.301	0.299	0.298	0.074	0.064	0.071
BH ₂ ⁺ (R_{st})	2.041	1.682	1.684	0.597	0.302	0.307
BeH ₂ (R_e)	0.282	0.265	0.278	0.080	0.049	0.077
BeH ₂ (R_{st})	2.695	1.883	1.892	0.875	0.275	0.293

Equilibrium distances (R_e) at experimental or optimized bond lengths and symmetrically stretched ones (R_{st}) at $R_{st} = 2.002 R_e$ (for LiH), $R_{st} = 2.676 R_e$ (for BeH⁺), $R_{st} = 1.599 R_e$ (for Li₂), $R_{st} = 1.487 R_e$ (for BH), $R_{st} = 1.826 R_e$ (for BH₂⁺), $R_{st} = 2.066 R_e$ (for BeH₂). Results for molecules correspond to standard STO-3G basis sets

four- and six-electron Be, LiH, BeH⁺, Li₂, BH, BH₂⁺, and BeH₂ and the basis sets STO-3G, in order to face up to an affordable computational cost. Moreover, we also report results corresponding to the Be atom in the cc-pVDZ basis set and the Mg one in the 6-31G basis sets, which are prototype examples of strongly correlated systems due to the near-degeneracies between its *s* and *p* shells. The molecular systems have been studied at equilibrium distances (R_e) and at stretched ones (R_{st}). The experimental geometrical distances have been used for the neutral species LiH, Li₂, BH, and BeH₂ [28]; in the molecular ion BeH⁺, we have used the internuclear distance reported in Refs. [29] and [30], while in the system BH₂⁺, the geometry was optimized with the GAUSSIAN package [31] at single and double excitations. The one- and two-electron integrals and the Hartree–Fock canonical molecular orbitals basis sets required for our calculations have been obtained from a modified version of the PSI 3.3 code [32]. We have constructed our own codes to determine the ground-state FCI wave functions for these systems expressed in the basis sets of CMO and NO; the orbitals minimizing the seniority number for a given wave function have been obtained from an iterative procedure reported in Ref. [27], using the CMO sets as initial bases of that iteration. The results found for the I_C and I_W quantities in those systems are gathered in Table 1, while Table 2 collects the I_Ω index values of each

seniority number level in the corresponding wave function expansion.

4 Discussion

The numerical results reported in Tables 1 and 2 have been obtained from the FCI method which coincides with the CI ($\Omega = 0, 2, 4$) one, in the CI framework, for the 4-electron systems (Be, LiH, and BeH⁺). Likewise, for the case of the 6-electron systems (Li₂, BH, BH₂⁺, and BeH₂), the FCI and the CI ($\Omega = 0, 2, 4, 6$) methods are identical. A survey of the results included in Table 1 shows that all described systems present low values for the I_W index, mainly at equilibrium geometries as well as at stretched ones in the NO and M_{\min} molecular basis sets. It means that most of the Slater determinants involved in expansion (3) can be grouped into a weight W_Ω , constituting a narrow Ω -level distribution. In fact, the weights corresponding to $\Omega = 0$ for these closed-shell singlet ground states are close to unity ($W_{(\Omega=0)} \sim 1$) [6]. Consequently, the determinants $\Lambda(\Omega = 0)$ are quite dominant in those expansions, while the others $\Lambda(\Omega \neq 0)$ can be neglected. The CI ($\Omega = 0$) method has also been called doubly occupied configuration interaction (DOCI) [33], since their *N*-electron wave functions are expanded on all possible $\Lambda(\Omega = 0)$ determinants.

Table 2 Calculated values of the quantities I_Ω (Eq. 8) for the ground states of atomic and molecular systems described by FCI expansions expressed in the canonical molecular orbitals (CMO), in the orbitalswhich minimize the seniority number (M_{\min}) and in the natural orbitals (NO)

System	CMO				M_{\min}				NO			
	$I_{\Omega=0}$	$I_{\Omega=2}$	$I_{\Omega=4}$	$I_{\Omega=6}$	$I_{\Omega=0}$	$I_{\Omega=2}$	$I_{\Omega=4}$	$I_{\Omega=6}$	$I_{\Omega=0}$	$I_{\Omega=2}$	$I_{\Omega=4}$	$I_{\Omega=6}$
Be(STO-3G)	0.648	2.585	0.522	–	0.648	2.115	2.008	–	0.648	2.837	2.000	–
Be(cc-pVDZ)	0.493	2.772	4.794	–	0.601	4.517	4.634	–	0.601	4.503	4.633	–
Mg(6-31G)	0.399	2.710	6.400	8.462	0.511	3.807	6.843	8.110	0.511	4.982	7.134	8.430
LiH(R_e)	0.132	1.944	2.194	–	0.166	3.120	2.018	–	0.166	3.155	2.017	–
LiH(R_{st})	0.807	2.063	2.207	–	0.862	3.604	2.234	–	0.862	3.608	2.234	–
BeH ⁺ (R_e)	0.132	1.956	2.124	–	0.163	3.405	2.007	–	0.163	3.414	2.007	–
BeH ⁺ (R_{st})	0.986	1.221	2.210	–	0.989	3.596	2.181	–	0.989	3.603	2.181	–
Li ₂ (R_e)	0.572	2.109	5.328	5.059	0.594	4.262	5.091	5.664	0.594	4.249	5.098	5.663
Li ₂ (R_{st})	0.822	2.033	5.261	5.027	0.849	4.633	5.080	6.347	0.849	4.639	5.079	6.346
BH(R_e)	0.456	2.712	3.226	0.604	0.542	1.949	3.235	2.393	0.539	1.991	3.251	3.645
BH(R_{st})	0.672	2.971	3.020	1.650	0.771	2.638	3.001	3.355	0.631	2.563	3.006	3.073
BH ₂ ⁺ (R_e)	0.210	3.977	2.160	2.923	2.213	3.259	2.383	1.856	0.210	3.493	2.168	2.665
BH ₂ ⁺ (R_{st})	1.324	2.627	2.242	0.677	1.331	3.031	2.255	2.897	1.327	3.014	2.252	2.245
BeH ₂ (R_e)	0.183	3.571	2.063	3.760	0.199	3.571	2.353	1.860	0.183	3.578	2.072	2.704
BeH ₂ (R_e)	1.635	2.981	2.178	3.210	1.572	3.343	2.254	0.625	1.560	3.310	2.250	3.136

Equilibrium distances (R_e) at experimental or optimized bond lengths and symmetrically stretched ones (R_{st}) at $R_{st} = 2.002 R_e$ (for LiH), $R_{st} = 2.676 R_e$ (for BeH⁺), $R_{st} = 1.599 R_e$ (for Li₂), $R_{st} = 1.487 R_e$ (for BH), $R_{st} = 1.826 R_e$ (for BH₂⁺), $R_{st} = 2.066 R_e$ (for BeH₂). Results for molecules correspond to standard STO-3G basis sets

The low values found for the I_W index show that, in these FCI wave functions, the CI ($\Omega = 0$) or DOCI expansions are close to the FCI ones, which is in agreement with the conclusions reported in Ref. [34] where the DOCI method has been picked up as a valuable tool to describe a wide variety of systems possessing strong correlation. As shown in Table 1, the I_C values are higher than their counterpart I_W ones, indicating that the expansions (3) involve significant contributions of $\Lambda(\Omega = 0)$ Slater determinants other than the ground closed-shell ones, which are usually chosen as reference determinants within the traditional excitation CI approach. This result is confirmed by the values reported in Table 2, where the $I_{\Omega=0}$ indices show a configurational distribution that cannot be considered as narrow. The results for the Be and Mg atoms show that these general trends are kept when basis sets larger than minimal STO-3G ones are used. The presence of strong correlation in the Be atom is well known, and consequently, its wave functions possess a multiconfigurational character even at zeroth-order descriptions; identical behavior has been found in the Mg atom. Our results confirm this feature showing that in the three isoelectronic species Be, LiH(R_e), and BeH⁺(R_e), the highest I_C index value corresponds to the Be atom (the widest multiconfigurational distribution), while the I_W index presents small values for that atomic system; its wave functions have a narrow distribution in terms of seniority levels, with low contribution of $\Lambda(\Omega \neq 0)$ determinants.

The results reported in Table 1 also allow one to compare, in terms of the values of the indices I_C and I_W , the expansions of the wave functions of these systems according to the molecular orbital basis sets in which they are expressed. As can be seen from that table, the values of both indices are considerably lower in the NO and M_{\min} basis sets than in their CMO counterparts (except for the Be atom in the STO-3G basis set); the Be atom recovers the improvement in the M_{\min} and NO molecular basis sets when the larger cc-pVDZ basis set is used. These results again confirm that the NO and M_{\min} molecular basis sets lead to more compact wave functions, as has been reported in Refs. [6, 9, 11]. These values also point out that the I_C and I_W indices constitute suitable devices to describe quantitatively the compactness of a wave function. The high values found for the I_C indices in the Be and Mg atoms in the three molecular basis sets can be interpreted in terms of the strong correlation exhibited by those systems. The appropriate ground-state wave functions for these atoms require several dominant Slater determinants. The I_Ω values reported in Table 2 reflect that seniority levels with very low contribution to the wave functions can present a broad determinantal distribution, i.e., the Li₂ molecule exhibits $I_{\Omega=4} > 5$ values because its $W_{\Omega=4} = 10^{-4}$ weight is expanded on 7560 Slater determinants in the STO-3G basis set [6]. Moreover, the $I_{\Omega=0}$ index values reported in that table indicate that all systems possess a narrower

distribution at equilibrium distances in the three molecular basis sets. Likewise, the molecular system descriptions at the stretched geometries systematically present higher values for the I_C and I_W indices (Table 1) than their counterparts at the equilibrium distances. This effect is interpreted in the framework of the progressive openness of the chemical bonds until their complete dissociation, which is reflected in the values of both indices. However, the index I_C shows a more sensitive character than the I_W one, and consequently, its use must be favored in order to account for the influence of the bond stretching on the wave function features.

5 Concluding remarks and perspectives

In this work, we have extended the formulation of the Shannon entropy indices, informational content (I_C), cumulative (I_W), and specific Ω -subspace (I_Ω), within the framework of the seniority number criterion for constructing N -electron wave function expansions in terms of Slater determinants. The quantitative evaluation of these indices has allowed us to implement analyses of the wave function expansions, determining their compactness in the well-known canonical molecular orbital and natural orbital basis sets, as well as in the recently proposed molecular orbital basis set which minimizes the seniority number of a given wave function. The results obtained for several atomic and molecular systems described at the FCI level show the suitability of the seniority-based formulation of these indices to measure quantitatively the wave function expansion compactness, as well as to analyze their multiconfigurational structure. We have also studied the ability of these indices to provide information on the evolution of the wave functions according to the stretching of the chemical bondings. We are currently working in our laboratories on the formulation of unitary transformations of molecular basis sets leading to the minimization of the Shannon entropy indices, in order to achieve a higher improvement on the compactness of wave function expansions.

Acknowledgments This work has been financially supported by the Grant Nos. GIU12/09 and UFI11/07 (Universidad del País Vasco, Spain), UBACYT 20020100100197 (Universidad de Buenos Aires, Argentina), PIP No. 11220090100061, 11220090100369, 11220130100377CO, and 11220130100311CO (Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina). We thank the Universidad del País Vasco for allocation of computational resources.

References

- Shavitt I (1998) *Mol Phys* 94:3
- Sherrill CD, Schaefer HF III (1999) *Adv Quantum Chem* 34:143 and references therein
- Bytautas L, Ivanic J, Ruedenberg K (2003) *J Chem Phys* 119:8217
- Giesbertz KJH (2014) *Chem Phys Lett* 591:220
- Bytautas L, Henderson TM, Jiménez-Hoyos CA, Ellis JK, Scuseria GE (2011) *J Chem Phys* 135:044119
- Alcoba DR, Torre A, Lain L, Massaccesi GE, Oña OB (2013) *J Chem Phys* 139:084103
- Stein T, Henderson TM, Scuseria GE (2014) *J Chem Phys* 140:214113
- Boguslawski K, Tecmer P, Limacher PA, Johnson PA, Ayers PW, Bultinck P, De Baerdemacker S, Van Neck D (2014) *J Chem Phys* 140:214114
- Alcoba DR, Torre A, Lain L, Massaccesi GE, Oña OB (2014) *J Chem Phys* 140:234103
- Limacher PA, Kim TD, Ayers PW, Johnson PA, De Baerdemacker S, Van Neck D, Bultinck P (2014) *Mol Phys* 5–6:853
- Alcoba DR, Torre A, Lain L, Oña OB, Capuzzi P, Van Raemdonck M, Bultinck P, Van Neck D (2014) *J Chem Phys* 141:244118
- Ring P, Schuck P (1980) *The nuclear many-body problem*. Springer, New York
- Koltun DS, Eisenberg JM (1988) *Quantum mechanics of many degrees of freedom*. Wiley, New York
- Garza AJ, Jiménez-Hoyos CA, Scuseria GE (2013) *J Chem Phys* 138:134102
- Jiménez-Hoyos CA, Rodríguez-Guzmán R, Scuseria GE (2013) *J Chem Phys* 139:204102
- Evangelista FA (2014) *J Chem Phys* 140:124114
- Mentel LM, Van Meer R, Gritsenko OV, Baerends EJ (2014) *J Chem Phys* 140:214105
- Ivanov VV, Lyakh DI, Adamowicz L (2005) *Mol Phys* 103:2131
- Kullback S (1959) *Information theory and statistics*. Wiley, New York
- Mathai AM, Tathie PN (1975) *Basic concepts in information theory and statistics*. Wiley, New York
- Pfeiffer PE (1978) *Concepts of probability theory*. Dover, New York
- Paldus J, Jeziorski B (1988) *Theor Chim Acta* 73:81
- Lain L, Torre A, Karwowski J, Valdemoro C (1988) *Phys Rev A* 38:2721
- Torre A, Lain L, Millan J (1993) *Phys Rev A* 47:923
- Lain L, Torre A (1995) *Phys Rev A* 52:2446
- Surjan PR (1989) *Second quantized approach to quantum chemistry*. Springer, Berlin
- Subotnik JE, Shao Y, Liang W, Head-Gordon M (2004) *J Chem Phys* 121:9220
- Johnson RD III (ed) (2006) *Computational chemistry comparison and benchmark database*. NIST Standard reference database vol 101. <http://www.srdata.nist.gov/cccbdb>
- Roos JB, Larson M, Larson A, Orel AE (2009) *Phys Rev A* 80:112501
- Chakrabarti K, Tennyson J (2012) *Eur Phys J* 66:31
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam MJ, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels

- AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, Revision D.01. Gaussian Inc., Wallingford
32. Crawford TD, Sherrill CD, Valeev EF, Fermann JT, King RA, Leininger ML, Brown ST, Janssen CL, Seidl ET, Kenny JP, Allen WD (2007) *J Comput Chem* 28:1610
33. Weinhold F, Wilson EB (1967) *J Chem Phys* 46:2752
34. Henderson TM, Bulik IW, Stein T, Scuseria GE (2014) *J Chem Phys* 141:244104