UNIVERSITEIT GENT

biblio.ugent.be

The UGent Institutional Repository is the electronic archiving and dissemination platform for all UGent research publications. Ghent University has implemented a mandate stipulating that all academic publications of UGent researchers should be deposited and archived in this repository. Except for items where current copyright restrictions apply, these papers are available in Open Access.

This item is the archived peer-reviewed author-version of:

Alternative Kullback-Leibler Information Entropy for Enantiomers.

Janssens, S.; Bultinck, P.; Borgoo, A.; Van Alsenoy, C.; Geerlings, P.

J. Phys. Chem. A, 2010, 114, 640–645.

http://pubs.acs.org/doi/abs/10.1021/jp9081883

To refer to or to cite this work, please use the citation to the published version:

Janssens, S.; Bultinck, P.; Borgoo, A.; Van Alsenoy, C.; Geerlings, P.; Alternative Kullback-Leibler Information Entropy for Enantiomers. J. Phys. Chem. A, 2010, 114, 640–645. http://pubs.acs.org/doi/abs/10.1021/jp9081883

Alternative Kullback-Leibler Information Entropy for Enantiomers

Sara Janssens[a+], Patrick Bultinck[b+], Alex Borgoo[a+], Christian Van Alsenoy[c], and Paul Geerlings[a+]∗

[a]Free University of Brussels (VUB), Faculteit Wetenschappen, Eenheid Algemene Chemie (ALGC), Pleinlaan 2, B-1050 Brussels, Belgium.

[b]Ghent University (UGent), Department of Inorganic and Physical Chemistry, Krijgslaan 281 (S-3), 9000 Ghent, Belgium.

[c]University of Antwerp (UA), Department of Chemistry, Universiteitsplein 1, B-2610 Antwerp, Belgium.

[+]Members of the QCMM Alliance Ghent-Brussels.

Author email address: sajansse@vub.ac.be

Received date

*Corresponding author: Tel: +32 2 6293314, Fax: +32 2 6293317, E-mail: pgeerlin@vub.ac.be

Abstract

In our series of studies on quantifying chirality, a new chirality measure is proposed in this work based on the Kullback-Leibler information entropy. The index computes the extra information that the shape function of one enantiomer carries over a normalized shape function of the racemate, while in our previous studies the shape functions of the R and S enantiomers were used considering one as reference for the other. Besides being mathematically more elegant (symmetric, positive definite, zero in the case of a non-chiral system), this new index bears a more direct relation with chirality oriented experimental measurements such as circular dichroism CD and optical rotation measurements, where the racemate is frequently used as a reference. The five chiral halomethanes holding one asymmetric carbon atom and H, F, Cl, Br and I as substituents have been analyzed. A comparison with our calculated optical rotation and with Avnir's Continuous Chirality Measure (CCM) is computed. The results show that with this index the emphasis lies on the differences between the non-coinciding substituents.

Keywords: Density Functional Theory, Information Theory, Kullback–Leibler, racemate, chirality, information deficiency, Carbó Similarity Index, Avnir's Continuous Chirality Measure, QSSA, halomethane, enantiomers, backbone alignment.

1. Introduction

The basic concept of entropy in information theory considers how much randomness resides in a signal or in a random event, or how much information is carried by a signal. Shannon[1] further developed and generalized in 1984 the statistical concepts of entropy, thereby laying the foundations of information theory. Shannon's accomplishments lie in revealing a universality of the entropy functional, by obtaining a functional representing the entropy or the uncertainty associated with any probability distribution. This generalized Shannon entropy is termed the Kullback-Leibler information deficiency, which measures the extra information that a given density carries over a prior density. The first application of this measure in the electronic structure theory is the determination of the exchange factor α in the self-interaction-corrected modification of the Slater $X\alpha$ method[2].

In our group information theory and the central role of entropy in it has been used as a tool to quantify chirality. The discrete black or white character of chirality has in the last decade been replaced by a continuously varying property, as promoted earlier by Avnir et al.[3, 4], Petitjean et al.[5, 6] and Fowler[7, 8], among others. This implies that one molecule can be more or less chiral compared to another molecule. In our previous articles a link is assumed between the degree of chirality and the (dis)similarity of two enantiomers, leading a quantification of the chirality to a quantification of the molecular similarity. Therefore quantum molecular similarity indices (QSI) based on the electron density were used, as presented – among others – by Carbó[9] and Hodgkin and Richards[10]. The

fundamental role of the electron density in the QSI connects our work with Density Functional Theory[11] in general and Conceptual DFT[12, 13, 14, 15] in particular.

In this earlier work we performed a similarity analysis applied to enantiomers of the halomethane CHFClBr[16], halogen substituted ethanes[17] and 1,3-disubstituted allenes[18], representing prototypical chiral systems with respectively one or two asymmetric centra or a chiral axis, using global and local Carbó similarity indices[9] (SI) with particular attention to a correlation with the optical rotation.

The next step in our chirality measure exploration based on the electron density was then to perform information theory[19] based calculations, which recently gained widespread interest in the quantum chemical community[20]. Gadre et al.[21, 22] for example examined a maximum-entropy procedure for several atomic systems by applying information theory techniques for the refinement of Gaussian basis sets. Nagy and Parr[23] argued that information deficiency gives a measure for the quality of an approximate electron wave function. The work of Borgoo et al.[24, 25] on the information deficiency of atoms, shows that an information theory based approach is markedly superior to quantum similarity calculations in revealing patterns of periodicity. An entity recently gaining more and more attention in quantum chemical literature[26, 27, 28], is the shape function $\sigma(\mathbf{r})$[29], defined as the density per particle $\rho(\mathbf{r})/N$. Ayers[26] and Bultinck et al.[30] illustrated, in an extension of the Hohenberg and Kohn theorems to the shape function, that this property is equivalent to the electron density $\rho(\mathbf{r})$ as carrier of information. When replacing in the work of Borgoo et al.[24] the density function by the shape function $\sigma(\mathbf{r})$, i.e. the density per particle $\rho(\mathbf{r})/N$, the results even improved, confirming the fundamental role of the shape function as information carrier[26, 27, 31, 32].

5

In the author's previous publication[33] information theory was used as a quantum chirality indicator. It was evaluated for chiral halomethanes possessing one asymmetric carbon atom and H, F, Cl, Br and I as substituents, resulting in five different enantiomeric pairs, and also for two halogen substituted ethanes.

In the present work we suggest an alternative chirality indicator that fulfils some elementary requirements of a distance[8]. Besides fulfilling these requirements, the new measure also has some similarity to what is usually done experimentally in e.g. circular dichroism experiments. When recording a spectrum, one must always consider a blank as a reference. This can either be the solvent alone, or a better blank is the spectrum of a racemic mixture. Likewise, we choose in our new measure the "racemic shape function" (see below) as a blank. It was analyzed for the same five halomethanes as used in our previous work, and evaluates the difference in information between the shape function of one enantiomer and a normalized shape function of the racemate.

Due to the dependence of the information deficiency on the relative orientation of the molecules, different alignments will be considered. The results will furthermore be compared to results of another chirality measure, the Continuous Chirality Measure (CCM) developed by Avnir et al.[3], and to results of the specific optical rotation.

2. Theoretical and Computational Details

2.1. Information Theory

It is only in the twentieth century that the concept of information entropy[19] acquired a precise mathematical description. Aczél and Daróczy[34] claimed that nearly all great inventions of civilization serve either to transform, store and transmit energy or to transform, store and transmit information. Also for quantum chemical research the interest in information theory recently developed widely.

This mathematical description of entropy is connected with missing information or uncertainty, measured through the Shannon entropy functional[1], which is defined as

$$S = -k \sum_i p_i \ln p_i \qquad (1)$$

where the sum runs over all events i with associated probability p_i and k is an arbitrary real constant. Naturally $p_i \geq 0$ and $\sum_i p_i = 1$. The Shannon entropy is maximal for a uniform distribution, which corresponds to spreading the probability distribution as evenly as possible.

A generalization of this Shannon entropy from discrete probability distributions to continuous ones may be expressed as

$$S = -\int p(x) \ln p(x) dx \qquad (2)$$

where S is a functional of the probability density function p.

Shannon noticed however[1] that the entropy is absolute in the discrete case, whereas in the continuous case it is relative to the coordinate system, implying it is not invariant to a transformation of variables. This invariance problem may be solved by introducing a reference probability density p_0. This measure is termed the Kullback-Leibler entropy[19, 35]. Omitting the negative sign one obtains equation (**3**).

$$\Delta S\left[p \,/\, p_0\right] = \int p(x)\ln\frac{p(x)}{p_0(x)}dx \qquad\qquad (\mathbf{3})$$

This quantity is often referred to as information deficiency, entropy deficiency, missing information or information entropy. It expresses the information discrimination or the information "distance" between the two distributions p and p_0. We may also say that this measure reveals the extra information that a given distribution p carries over the prior distribution p_0. Notice that ΔS equals zero when $p(x) = p_0(x)$; when $p(x)$ and $p_0(x)$ are equally normalized one can prove that $\Delta S \geq 0$ by employing the inequality $\ln(1\,/\,x) \geq (1-x)$ (for positive values of x)[36].

Based on these considerations, we propose the following expression as a quantification of the similarity of two enantiomers R and S of a chiral molecule with electron distributions $\rho_R(\mathbf{r})$ and $\rho_S(\mathbf{r})$ and shape functions $\sigma_R(\mathbf{r})$ and $\sigma_S(\mathbf{r})$:

$$\Delta S_{racemate} = \int [\sigma_R(\mathbf{r}) - \sigma_{RS}(\mathbf{r})] \ln \frac{\sigma_R(\mathbf{r})}{\sigma_{RS}(\mathbf{r})} d\mathbf{r} \qquad (4)$$

with as shape function for the racemate *RS*

$$\sigma_{RS}(\mathbf{r}) = \frac{\rho_R(\mathbf{r}) + \rho_S(\mathbf{r})}{2N} = \frac{\sigma_R(\mathbf{r}) + \sigma_S(\mathbf{r})}{2} \qquad (5)$$

This property can be interpreted as a measure for the information divergence between the shape functions of the *R* enantiomer and the racemate *RS*, or the information in *R* distinguishing it from *RS*, or also the extra information that the *R* enantiomer carries over the racemate *RS*. We use shape functions $\sigma(\mathbf{r})$ instead of electron densities $\rho(\mathbf{r})$ because these are equally normalized, making the expression more generally applicable (vide supra).

This expression (**4**) has some noteworthy mathematical features, promoting it as an interesting alternative to our earlier information deficiency.

The integrand is positive definite. If $\sigma_R(\mathbf{r}) < \sigma_{RS}(\mathbf{r})$, the first term is negative but also the second one, making the result still positive. As a consequence of this positive definite integrand, also the proposed integral in expression (**4**) is positive definite, which can easily be shown by using the inequality $\ln(x) \leq x - 1$ and taking into account that any shape function involved normalizes to 1, as required for a shape function[29]. The integral may be divided into two separate integrals of the same type as equation (**3**), both of which can be shown to be positive. When looking at non-chiral molecules,

yielding $\sigma_R(\mathbf{r}) = \sigma_{RS}(\mathbf{r})$, the integral is zero, thus resulting in a chiral information loss equal to zero.

Moreover, the expression is symmetric, in other words it satisfies the equality $\Delta S\left[\sigma_R / \sigma_S\right] = \Delta S\left[\sigma_S / \sigma_R\right]$. Despite being positive definite and symmetrical, our measure can not be considered a true distance as it can not be shown to satisfy the triangular inequality, stating that for any three points $x, y, z \in X$ of the metric space (X, d), we have

$$d(x,y) \leq d(x,z) + d(z,y) \qquad\qquad (6)$$

Next to these mathematical advantages, this equation also has an important chemical benefit. We use the racemic mixture – a mixture that has equal amounts of left- and right-handed enantiomers of the chiral molecule – as a reference for one enantiomer. This is comparable to experimental measurements of circular dichroism and optical rotation, where the racemate is used as a blanco reference solution. This reference spectrum or optical rotation angle is then subtracted from the "brute" values so that background noise may be reduced.

This newly proposed measure of chirality in equation (**4**) can be compared to our previous information entropy measure (**3**), which did not possess the direct link with the experiment but which had been investigated for atomic densities before[24, 25, 24]. Although the integral of our previous measure is positive definite, the integrand is not always

positive. Equation (**4**) is – according to our knowledge – the closest we can get to a distance with information entropy[37].

We will also try to relate the results to calculations of Avnir's Continuous Chirality Measure (CCM)[3] as an alternative, based on the comparison of the chiral structure and its closest non-chiral compound. Petitjean[38] mentioned two methods to compare enantiomers, namely comparing one enantiomer to the other one as in our chirality measures, or comparing the enantiomers to an achiral reference as in Avnir's CCM. In the context of Avnir's CCM the chirality content of a given set of points $\{P_i\}_{i=1}^n$ is determined from its distance to the nearest configuration of points $\{\hat{P}_i\}_{i=1}^n$ which has an improper element of symmetry. The distance between the two sets of points is defined by

$$\mathbf{S'}(G_{achiral}) = \frac{1}{n} \sum_{i=1}^n \left\| P_i - \hat{P}_i \right\|^2 \tag{7}$$

When applied to molecules the set of points $\{P_i\}_{i=1}^n$ can be taken as the position of the different atoms in space.

The $S'(G_{achiral})$ thus obtained is the minimal chirality measure of the given configuration and lies due to scaling[3] on a continuous range of $0 \le S' \le 1$ or $0 \le S \le 100$ with $S = 100 S'$.

A link between the dissimilarity or information deficiency and the optical rotation of enantiomers, as was previously put forward by Mezey et al.[39], might not be unexpected

11

since the optical rotation – quantified by the specific rotation $[\alpha]_D$ – is an experimentally accessible quantity for the degree of chirality of a molecule and since one might assume a description of the chirality by the dissimilarity between enantiomers. Note however that previous work[17, 33, 40] has shown that prudence needs to be taken when stating a positive correlation between the global similarity or information deficiency and the optical rotation.

The specific rotation $[\alpha]_\nu$[41] of a chiral molecule, with molecular weight M, in dilute solution measured at frequency ν is given by

$$[\alpha]_\nu = \frac{28800\pi^2 N_{Av}\nu^2}{c^2 M}\gamma_S \beta(\nu) \qquad (8)$$

where $\beta = \frac{1}{3}Tr\left[\beta_{\alpha\beta}\right]$, with $\beta_{\alpha\beta}$ the electric dipole – magnetic dipole polarisability tensor; γ_s is the condensed phase solvent effect. The choice for the frequency ν in this work is the traditional so-called D-line of Sodium.

2.2. Alignment of the Enantiomers

An important complication of the molecular information deficiency is the dependence on the relative orientation of both enantiomers. Several approaches have been proposed to resolve this problem. Some examples of alignment possibilities are alignments according to common physico-chemical features, or topological and geometrical features (TGSA[42]), or aligning the molecules so that the resulting molecular similarity is maximized

(QSSA[43, 44, 45, 46, 47]). It deserves mentioning that for a true distance, one can demand internal consistency over a molecular set[46]. As the information measures do not resemble true distances in all aspects, the dependence of the resulting index on the molecular alignment may remain significant.

The approach followed in this paper is the so-called backbone (BB) alignment, already proven in previous papers to be an efficient alignment method. This choice gives us the possibility – as opposed to TGSA – to evaluate global as well as local entropy measures, for the asymmetric carbon atoms and the mono-atomic (non-asymmetric) substituents. In the case of halomethanes, we superimpose the asymmetric carbon atom and two of its substituents, yielding 6 different alignments for each pair of enantiomers. This means that in equation (**4**) six different ways are considered to align the *R* and the *S* enantiomer of the racemate *RS*, while the separate *R* enantiomer fully coincides with the *R* enantiomer of the racemate.

This alignment problem does not occur when using CCM, generating only one value for each enantiomeric pair.

2.3. Computational Details

All geometry optimizations, electron densities and optical rotations used in this work were calculated in a DFT approach using the GAUSSIAN 03 program[48] at the B3LYP/6-311G* level[49, 50]. We implemented the necessary integrals needed to evaluate the information deficiency in the BRABO program package developed by Van Alsenoy et

al.[51, 52], already containing the implementation of global and local Carbó similarity indices.

Optical rotations have been calculated using ab initio density functional theory with gauge-invariant atomic orbitals (GIAOs).

The CCM values[3] were calculated using Avnir's freely available program at http://www.csm.huji.ac.il.

3. Results and discussion

In this section we show and interpret the results of the proposed information deficiency measure, with special attention for its capability in quantifying the chirality of halomethanes. These molecules were chosen for their prototypical character as chiral systems with a single asymmetric carbon atom. We include chiral halomethanes with H, F, Cl, Br and I as substituents, providing five different enantiomeric pairs. Following the backbone alignment a superposition of the asymmetric carbon atom and two of its substituents is made, yielding six different alignments for each pair of enantiomers. The separate *R* enantiomer has the same alignment as the *R* enantiomer of the racemate *RS*, so both *R* molecules are fully coinciding and have six different alignments with the *S* enantiomer of the racemate.

A comparison of this new information deficiency with our previous Kullback-Leibler entropy deficiency and Avnir's Continuous Chirality Measures will be discussed and related to optical rotation values.

3.1. Information Deficiency

The results of the information deficiency with the racemate as a reference, equation (**4**), is shown in **Figure 1** for the six different alignments of all five enantiomeric pairs. Each alignment is represented by the sum of the nuclear charges of the coinciding atoms. This gives us the possibility to differentiate between the different alignments of an enantiomeric pair. In this case, the sum of the nuclear charges of the coinciding atoms can be considered as proportional to the sum of the substituent volumes, whose diversity might be expected to be related to the chirality[53].

We suggest focusing on the difference of the nuclear charges ΔZ of both non-coinciding atoms for each alignment. For example the alignment of CHFClBr with coinciding atoms CHF, providing a value of 16 as its abscissa (sum of the nuclear charges of the coinciding atoms) on graph **Figure 1**, has the atoms Cl and Br non-coinciding resulting in a value $\Delta Z = Z_{Br} - Z_{Cl}$ of 18. Considering these ΔZ values for the six different alignments of molecule CHFClBr shows a trend between this value and the information deficiency. The entropy deficiency becomes smaller, and thus the similarity between both enantiomers bigger, as the value of ΔZ for the non-coinciding atoms gets smaller, or the non-coinciding substituents become more similar. We observe a first ordering for the heaviest non-coinciding atom, and a second ordering for the lighter non-coinciding atom. A similar tendency is obtained for all five considered enantiomeric pairs.

Notice that the smallest information deficiency value is reached for the alignment with the heaviest substituents coinciding (so coinciding atoms C*, Cl and Br for CHFClBr in **Figure 1**). This indicates that the information distinguishing the *R* enantiomer from the racemate *RS*, aligned this way, is small. We may also say that the *R* enantiomer contains a lot of information about the racemate *RS*. Thus a small value of the information deficiency leads to the most similar alignment for the enantiomer and the racemate. With our previous information deficiency measure[33] the same trend was found for the two enantiomers.

The largest entropy is found when the lightest and heaviest substituents do not coincide (e.g. coinciding atoms C*, F and Cl; non-coinciding atoms H and Br for CHFClBr in **Figure 1**). With the new information deficiency actually three molecules are involved in the alignment, namely the separate *R* enantiomer and the *R* and the *S* enantiomer of the

racemate *RS*. Both *R* enantiomers are fully coinciding. This means that in the alignment where the two heaviest substituents do not coincide in the *R* and the *S* enantiomer (e.g. coinciding atoms C*, H and F; non-coinciding atoms Cl and Br for CHFClBr in **Figure 1**), they still do coincide in both *R* enantiomers. Thus this alignment with the two heaviest substituents not coinciding will in this case not result in the largest entropy value, as was the case for our previous information deficiency.

3.2. Carbó's quantum similarity index

These new information deficiency results in **Figure 1** can be compared to Carbó's global quantum similarity results, which may be found in **Figure 2** for the CHFClBr halomethane. In **Figure 2** we differentiate two subsets of values, three very high similarity values for alignments where the bromine atoms are coinciding, and three very low values when these heavy atoms are not coinciding. Within both subsets the values differ little. The SI is most sensitive to the relative position of the heavy atoms, and is almost indifferent to that of the lighter atoms.

The results in **Figure 1** may be compared to these observations. High similarity values correspond very nicely to low information deficiency values for coinciding atoms. Notice that the information deficiency evolves more gradually, making it more appealing than the QSI as the basis of a chirality measure.

3.3. Single value for each molecule

In this work we aimed for one single value characterizing the chirality of a molecule, in other words, a value independent of the alignment of both enantiomers. To achieve a single value for each molecule, different options are possible.

Initially we looked at the arithmetic mean of the information deficiency of the six alignments for each molecule, as a function of the sum of the nuclear charges of all atoms. From the data in **Table 1** a linear correlation with a negative slope may be deduced when the four molecules containing I are considered, indicating that the heavier enantiomeric pairs are more similar compared to the lighter enantiomeric pairs.

A second solution is to consider only the alignment resulting in the highest information deficiency value, the best score alignment. This is the alignment where the heaviest and the lightest substituents do not coincide (e.g. H and Br not coinciding, so C*, F and Cl coinciding for the CHFClBr molecule). This highest entropy alignment shows a similar tendency as the mean value of the six alignments.

Another possibility to obtain a single value for each molecule would be to use the Quantum Similarity Superposition Algorithm (QSSA), aligning the molecules so that the resulting molecular similarity is maximized, instead of the Backbone alignment. A closer look at this alignment reveals that this maximal similarity value is obtained when the three heaviest atoms of both enantiomers coincide; so in the case of the CHFClBr molecule the atoms F, Cl and Br will coincide. The reason for this behavior of QSSA is that the overlap quantum similarity measure is highly dependent on overlapping the highest density areas, irrespective of the nature of the atoms. In the case of our new information deficiency measure, where the racemate is used as a reference, we do not find this alignment between the racemate and one enantiomer chemically relevant.

Yet another possibility we considered was to use the closest non-chiral structure to the enantiomer, as used in Avnir's CCM. Also with this non-chiral structure as a reference for one enantiomer, a clear-cut chemically relevant alignment between the non-chiral structure and the enantiomer does not seem possible.

3.4. Optical Rotation

In **Figure 3** the result of the absolute value of the specific optical rotation $[\alpha]_D$ for the five different halomethanes is plotted against the average information deficiency. No clear trend or empirical link is found between the information deficiency and the optical rotation. This shows that the optical rotation, which is known to be not exclusively related to geometrical changes, is even with the addition of electronical changes not easily recovered. Also the overall values of the optical rotation are very small, making them not really informative and maybe not well suited for a correlation with information deficiency values. It must be stressed that optical rotation values are strongly dependent on the chosen frequency ν. In this work we opted for the traditional so-called D-line of Sodium.

Note that also when looking at Avnir's CCM values, no correlation can be found with the optical rotation for our halomethanes. This is probably a consequence of the fact that the CCM relies only on nuclear locations, and not on electron densities. It has previously been shown[54] for the homologous series of [6]-helicene to [14]-helicene racemates that the CCM correlates nicely with the optical rotation as the electronic structures of the homologous series are very well represented by the nuclear coordinates.

3.5. Continuous Chirality Measure CCM

As already mentioned in paragraph 2.1., also the very popular chirality measure of Avnir et al.[3] has been calculated for our molecules. The first step in this approach is to find the nearest configuration of points $\left\{\hat{P}_i\right\}_{i=1}^{n}$ which has an improper element of symmetry, in other words the closest achiral structure. The tool used therefore is termed the *folding/unfolding procedure*, and is based on the construction of a symmetric object.

For the halomethane molecules, the closest achiral structure is a structure where all atoms are located in a plane with the substituents surrounding the carbon atom.

The results for the considered halomethanes are shown in **Table 1**, where it can be seen that our five enantiomeric pairs are very closely related, with all CCM **S** values lying between 17.40 and 22.82 on a scale of 0 to 100.

If comparing the CCM values with the mean information deficiency measure, a link or trend between both can hardly be observed. This might be expected as the CCM method relies only on nuclear locations, and not on electron densities as the information deficiency does. In contrast to Avnir's CCM, the advantage of information theory is that searching for the closest achiral structure is not required. The orientation issue occurring to information theory using the backbone alignment can easily be overcome by choosing the mean alignment or the best score alignment.

4. Conclusions

In this work we reported the calculation of an information theory based quantum similarity indicator, the Kullback-Leibler information deficiency using the racemate as a reference. This chirality indicator highlights how much information distinguishes one enantiomer (the *R* or the *S* enantiomer) from the racemate *RS*.

This measure was evaluated for enantiomers possessing a single asymmetric carbon atom, namely five chiral halomethane pairs with as substituents H, F, Cl, Br and I, each of these being superposed in six different alignments.

We find that the information deficiency between both enantiomers becomes smaller, or the similarity larger, when the difference of the nuclear charges of both non-coinciding substituents gets smaller, indicating that the non-coinciding atoms are more similar. This implies that using this index, the emphasis lies on the differences between the non-coinciding substituents.

With our previously published information deficiency measure, where one enantiomer is used as reference for the other one, the sum of the nuclear charges of the coinciding atoms is stressed. It showed smaller entropy values when heavier atoms coincided.

The results of the new information deficiency measure correlate well with Carbó's quantum similarity index. High similarity values correspond with low entropy deficiencies for coinciding atoms.

To obtain a single value characterizing the chirality of a molecule, we might take the arithmetic mean of the entropy values of the six different alignments, or the best score alignment.

Comparison with the optical rotation or Avnir's CCM does not show a clear trend, showing that even adding electronic structure aspects to a geometry based approach is still not enough to capture this intricate property in a non-homologous series.

An analysis of the integrand of expression (**4**) as a local measure, with as prior densities as well the racemate as each enantiomer separately, might reveal some interesting information about how the Kullback-Leibler information deficiency continuously describes chirality. This issue might be addressed in future work.

Acknowledgments

References:

(1)   Shannon, C.E. *Bell Syst. Tech. J.* **1948**, *27*, 379.

(2)   Raju, K. B. K. ; Nair, P. S. V. ; Sen, K. D. *Chem. Phys. Lett*., **1990**, *170*, 89.

(3)   Zabrodsky, H. ; Avnir, D. *J. Am. Chem. Soc.* **1995**, *117*, 462.

(4)   Alvarez, S. ; Alemany, P. ; Avnir, D. *Chem. Soc. Rev.* **2005**, *34*, 313 and references cited therein.

(5)   Petitjean, M. *C. R. Acad. Sci. Paris, Chimie/Chemistry* **2001**, 4, 331.

(6)   Petitjean, M. *Entropy*, **2003**, 5,271.

(7)   Rassat, A. ; Fowler, P.W. *Chem. – A Eur. J.* **2004**, *10*, 6575.

(8)   Fowler, P.W. *Symmetry: Culture and Science*, **2005**, *16*, 321.

(9)   Carbó, R. ; Arnau, M. ; Leyda, L. *Int. J. Quant. Chem.* **1980**, *17*, 1185.

(10)   Hodgkin, E. E. ; Richards, W. G. *Int. J. Quant. Chem.* **1987**, *14*, 105.

(11)   Parr, R.G.; Yang, W. In *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, Clarendon Press, Oxford, 1989.

(12)   Parr, R.G.; Yang, W. *Annu. Rev. Phys. Chem.* **1995**, *46*, 701.

(13)   Chermette, H. *J. Comp. Chem.* **1999**, *20*, 129.

(14)   Geerlings, P.; De Proft, F.; Langenaeker, W. *Chem. Rev.* **2003**, *103*, 1793.

(15) Geerlings, P.; Boon, G.; Van Alsenoy, C.; De Proft, F. *Int. J. Quant. Chem.* **2005**, *101*, 722.

(16) Boon, G.; Van Alsenoy, C.; De Proft, F.; Bultinck, P.; Geerlings, P. *J. Phys. Chem. A* **2003**, *107*, 11120.

(17) Janssens, S.; Van Alsenoy, C.; Geerlings, P. *J. Phys. Chem. A* **2007**, *111*, 3143-3151.

(18) Janssens, S.; Boon, G.; Geerlings, P. *J. Phys. Chem. A* **2006**, *110*, 9267.

(19) Kullback, S.; Leibler, R.A. *Ann. Math. Stat.* **1951**, *22*, 79.

(20) For a review, see Gadre, S. In *Review of Modern Quantum Chemistry*, Editor Sen K.D., World Scientific, Singapore, Volume 1, 2002, 108.

(21) Gadre, S.R.; Kulkarni, S.A.; Shrivastava, I.H. *Chem. Phys. Lett.* **1990**, *166*, 445.

(22) Gadre, S.R.; Sears, S.B.; Chakravorty, S.J.; Bendale, R.D. *Phys. Rev. A* **1985**, *32*, 2602.

(23) Nagy, A.; Parr, R.G. *Int. J. Quant. Chem.* **1996**, *58*, 323.

(24) Borgoo, A.; Godefroid, M.; Sen, K.D.; De Proft, F.; Geerlings, P. *Chem. Phys. Lett.* **2004**, *399*, 363.

(25) Borgoo, A.; Godefroid, M.; Indelicato, P.; De Proft, F.; Geerlings, P. *J. Chem. Phys.* **2007**, *126*, 44102.

(26) Ayers, P.W. *Proc. Natl. Acad. Sci. (USA)* **2000**, *97*, 1959.

(27) De Proft, F.; Ayers, P.W.; Sen, K.D.; Geerlings, P. *J. Chem. Phys.* **2004**, *120*, 9969.

(28) Ayers, P.W. ; De Proft, F. ; Geerlings, P. *Phys. Rev. A* **2007**, *75*, 12508.

(29) Parr, R. G.; Bartolotti, L. J. *J. Phys. Chem.* **1983**, *87*, 2810.

(30) Bultinck, P.; Carbó-Dorca, R. *J. Math. Chem.*, **2004**, *36*, 191.

(31) De Proft, F.; Geerlings, P. *J. Phys. Chem. A* **1997**, *101*, 5344.

(32) Sen, K. D.; De Proft, F.; Borgoo, A.; Geerlings, P. *Chem. Phys. Lett.* **2005**, *410*, 70.

(33) Janssens, S. ; Borgoo, A. ; Van Alsenoy, C. ; Geerlings, P. *J. Phys. Chem. A* **2008**, *112*, 10560.

(34) Aczél, J.; Daróczy, Z. In *On Measures of Information and Their Characterization*, Academic Press, New York, 1975.

(35) Kullback, S. In *Information Theory and Statistics* , Wiley, New York, 1959.

(36) Abramowitz, M.; Segun, I.A. In *Handbook of Mathematical Functions*, Dover publications, New York, 1965, 68.

(37) Ayers, P.W. *Theoretical Chem. Accounts* **2006**, *115*, 370.

(38) Petitjean, M. *Symmetry: Culture and Science*, **2006**, *17*, 197.

(39) Mezey, P. G. ; Ponec, R. ; Amat, L. ; Carbó-Dorca, R. *Enantiomer*, **1999**, *4,* 371.

(40) Boon, G. ; Van Alsenoy, C. ; De Proft, F. ; Bultinck, P. ; Geerlings, P. *J. Phys. Chem. A* **2006**, *110*, 5114.

(41) Stephens, P. J. ; Devlin, F. J. ; Cheeseman, J. R. ; Frisch, M. J. ; Mennucci, B. ; Tomasi, J. *Tetrahedron: Asymmetry*, **2000**, *11*, 2443.

(42) Gironés, X. ; Robert, D. ; Carbó-Dorca, R. *J. Comput. Chem.* **2001**, *22,* 255.

(43) Constans, P. ; Carbó-Dorca, R. *J. Chem. Inf. Comput. Sc.* **1995**, *35*, 1046.

(44) Constans, P. ; Amat, L. ; Carbó-Dorca, R. *J. Comput. Chem.* **1997**, *18,* 826.

(45) Amat, L. ; Carbó-Dorca, R. *Int. J. Quant. Chem.* **2002**, *87*, 59.

(46) Bultinck, P. ; Carbó-Dorca, R. ; Van Alsenoy, C. *J. Chem. Inf. Comput. Sc.* **2003**, *43*, 1208.

(47) Bultinck, P. ; Kuppens, T. ; Gironés, X. ; Carbó-Dorca, R. *J. Chem. Inf. Comput. Sc.* **2003**, *43*, 1143.

(48) Gaussian 03, Revision B.03, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.;

Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Pittsburgh PA, 2003.

(49) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(50) Lee, C. ; Yang, W. ; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(51) Rousseau, B. ; Peeters, A. ; Van Alsenoy, C. *Chem. Phys. Lett.*, **2000**, *324*, 189.

(52) Van Alsenoy, C. ; Peeters, A. J. Mol. Struct. (Theochem) 1993, 105, 19.

(53) Meyer, A.Y.; Avnir, D. Struct. Chem. 1991, 2, 475.

(54) Katzenelson, O. ; Edelstein, J. ; Avnir, D. Tetrahedron: Asymmetry, 2000, 11, 2695.

Tables:

Table 1: Arithmetic mean of the six alignments of the information deficiency $\Delta S_{racemate}$, and Avnir's Continuous Chirality Measure S, compared with the sum of the nuclear charges of all atoms for all five halomethanes.

halomethane	sum Z	mean $\Delta S_{racemate}$	CCM S
CHFClBr	68	0.813	19.47
CHFClI	86	1.176	18.19
CHFBrI	104	1.100	17.40
CHClBrI	112	0.933	18.16
CFClBrI	120	0.691	22.82

Figure captions:

Figure 1: Values of information deficiency $\Delta S_{\text{racemate}}$, as function of the six different alignments for all five enantiomeric pairs, represented by the sum of the nuclear charges of their coinciding atoms.

Figure 2: Global similarity index (using the total density) for the six different alignments of the halomethane CHFClBr, represented by the sum of the nuclear charges of their coinciding atoms.

Figure 3: The mean over all alignments of the information deficiency $\Delta S_{\text{racemate}}$ compared with theoretically calculated values of the specific optical rotation $[|\alpha|]_D$ for all five halomethanes.