Monte Carlo Approach to Total π-Electron Energy of Conjugated Hydrocarbons

I. Gutman and M. Rašković
Faculty of Science, University of Kragujevac

Z. Naturforsch. **40 a,** 1059–1061 (1985); received July 3, 1985

Approximating the π-electron molecular orbital energy levels by uniformly distributed random variables, a McClelland-type formula for the total π-electron energy is obtained. Conditions are determined under which a given distribution function will result in a formula of McClelland type.

The total π-electron energy (E), as calculated within the Hückel molecular orbital model, has for a long time been attracting the attention of theoretical chemists [1, 2]. Especially important seems to be the recent finding [2] that a precise linear relationship exists between E and the kinetic energy of the π-electrons.

In the following the conjugated molecule under consideration will be represented by its molecular graph [3]. This graph has n vertices and m edges. For the sake of simplicity we shall assume that the molecule considered is an alternant hydrocarbon and that n (= the number of carbon atoms) is an even number. If $x_1, x_2, \ldots, x_n$ are the eigenvalues [3] of the molecular graph, labeled in a non-increasing order, then

$$E = 2 \sum_{j=1}^{n/2} x_j . \tag{1}$$

A relation which will be needed in the forthcoming discussion is [3]

$$\sum_{j=1}^{n/2} x_j^2 = m . \tag{2}$$

Now, if $x_1, x_2, \ldots, x_{n/2}$ are viewed as random numbers, that is $n/2$ independent realizations of a random variable X, then their sample mean $e(X)$ and sample dispersion $d(X)$ are given by

$$e(X) = \frac{1}{n/2} \sum_{j=1}^{n/2} x_j = E/n , \tag{3}$$

$$d(X) = \frac{1}{n/2} \sum_{j=1}^{n/2} [x_j - e(X)]^2 = 2m/n - (E/n)^2 . \tag{4}$$

Since obviously $d(X) \geqq 0$, we get from (4)

$$E \leqq \sqrt{2mn} , \tag{5}$$

which is just the McClelland upper bound for the total π-electron energy [4].

In the same paper [4] in which the upper bound (5) has been deduced, it has been established that E can be fairly well reproduced by the expression

$$E = a \sqrt{2mn} , \tag{6}$$

where a is an empirical parameter, whose value is about 0.9. The accuracy of (6) was verified in [5], where a comparative study of a great number of approximate topological formulas has been performed.

It has been shown [6] that the upper bound (5) is related to the assumption that all the (positive) graph eigenvalues are degenerate. Namely, the following result was deduced in [6]:

Lemma 1. Under the constraint (2), the greatest possible value of E is $(2mn)^{1/2}$, and this maximum is reached if and only if $x_1 = x_2 = \cdots = x_{n/2}$.

On the basis of Lemma 1 one may be inclined to expect that the validity of the approximation (6) is also stipulated by a relatively small deviation of graph eigenvalues from their mean value. Surprisingly, however, if the opposite is true, namely if all eigenvalues are mutually distinct, an expression for E, closely similar to (6) exists [6]:

Lemma 2. If the numbers x_j are equidistant, that is $x_{j+1} - x_j = \text{const}$ for all $j = 1, 2, \ldots, n/2 - 1$, then under the constraint (2),

$$E = a(n) \sqrt{2mn} , \tag{7}$$

where the function $a(n) = [3n^2/(4n^2 - 4)]^{1/2}$ rapidly converges to $(3/4)^{1/2} = 0.87$ when n tends to infinity.

Reprint requests to Prof. Dr. Ivan Gutman, Faculty of Science, P.O. Box 60, YU-34000 Kragujevac, Yugoslavia.

0340-4811 / 85 / 1000-1059 $ 01.30/0. – Please order a reprint rather than making your own copy.

Comparing Lemmas 1 and 2 we see that quite different distributions of the graph eigenvalues may lead to McClelland-type formulas for the total π-electron energy. In the present paper further arguments in favour of this conclusion will be given. For the beginning we formulate the following result.

Lemma 3. If $x_1, x_2, \ldots, x_{n/2}$ are random numbers, uniformly distributed in the interval $(0, p)$, then for an arbitrary positive value of p,

$$\langle E \rangle = b\sqrt{2\langle m \rangle n}\,, \tag{8}$$

where $b = (3/4)^{1/2} = 0.87$ equals the limit value of the function $a(n)$ in (7). In (8), $\langle E \rangle$ and $\langle m \rangle$ denote the expectation values of E and m, as defined via (1) and (2).

Numerical work

In order to test Lemma 3 we have chosen the interval $(0, p)$ by generating a random number $r \in (0, 1)$ and setting $p = 1/r$. Then by generating the random numbers $r_j \in (0, 1)$, $j = 0, 1, 2, \ldots$ we defined $n/2 = [50\, r_0]$ and $x_j = p\, r_j$, $j = 1, 2, \ldots, n/2$ and calculated E and m by means of (1) and (2). After repeating this procedure 20 times (for a fixed value of p), we determined the parameter b in (8) by least squares fitting.

In Table 1 are presented ten typical results obtained in this manner.

Proof of Lemma 3

Suppose that X is a continuous random variable, uniformly distributed in the interval $(0, p)$. Then its distribution function $\varrho(x)$ is equal to $1/p$. Bearing in mind (1) and (2), we have

$$\langle E \rangle = (n/2)\int_0^p 2x\,\varrho(x)\,\mathrm{d}x = np/2\,, \tag{9}$$

$$\langle m \rangle = (n/2)\int_0^p x^2\varrho(x)\,\mathrm{d}x = np^2/6\,. \tag{10}$$

Elimination of the parameter p from (9) and (10) results immediately in (8). Note the very good agreement between the parameter b in (8) and the empirically determined value for a in McClelland's formula (6).

A more careful examination of the above proof shows that the assumption of the uniform distribution of the variable X is not at all necessary for the derivation of a McClelland-type formula for E.

Indeed, let X be a continuous random variable whose values lie in the interval $(0, p)$ and whose distribution function is $\varrho(x)$. Suppose that the first and the second moment of X are proportional to p and p^2, respectively:

$$\int_0^p x\,\varrho(x)\,\mathrm{d}x = c_1 p\,, \tag{11}$$

$$\int_0^p x^2\varrho(x)\,\mathrm{d}x = c_2 p^2\,, \tag{12}$$

where c_1 and c_2 are constants, independent of p. Then, according to (9) and (10),

$$\langle E \rangle = n\,c_1 p \tag{13}$$

and

$$\langle m \rangle = n\,c_2 p^2/2\,, \tag{14}$$

and therefore

$$\langle E \rangle = c\sqrt{2\langle m \rangle n}\,, \tag{15}$$

where

$$c = c_1 c_2^{-1/2}\,. \tag{16}$$

Clearly, the distribution function $\varrho(x)$ must depend in a pertinent way on the parameter p.

We show now that for an arbitrary function $f(x)$ which is non-negative in the interval $(0, 1)$ and whose integral

$$I = \int_0^1 f(x)\,\mathrm{d}x \tag{17}$$

Table 1. Monte Carlo estimation of the parameter b in (8). Each result is obtained by a least squares fitting of data for 20 "random molecules".

No.	b	Correlation coefficient
1	0.888	0.9995
2	0.879	0.99990
3	0.863	0.99990
4	0.854	0.998
5	0.862	0.9997
6	0.904	0.998
7	0.828	0.9998
8	0.855	0.9990
9	0.835	0.9997
10	0.882	0.998

exists, we can choose the distribution function $\varrho(x)$ as

$$\varrho(x) = (Ip)^{-1} f(x/p) , \tag{18}$$

and $\varrho(x)$ has the properties (11) and (12). As a matter of fact, for a distribution function of the form (18), the k-th moment of X is proportional to p^k. This can be easily seen by direct calculation of the k-th moment:

$$\int_0^p x^k \varrho(x)\,\mathrm{d}x = (p^k/I) \int_0^1 y^k f(y)\,\mathrm{d}y , \tag{19}$$

where $y = x/p$. Hence

$$c_1 = \int_0^1 x f(x)\,\mathrm{d}x \Big/ \int_0^1 f(x)\,\mathrm{d}x , \tag{20}$$

$$c_2 = \int_0^1 x^2 f(x)\,\mathrm{d}x \Big/ \int_0^1 f(x)\,\mathrm{d}x ,$$

and we have arrived at the following result:

Theorem 1. Let $f(x)$ be a function having non-negative values in the interval (0, 1), whose integral (17) exists. If $x_1, x_2, \ldots, x_{n/2}$ are random numbers, distributed in the interval $(0, p)$ according to (18), then for an arbitrary positive value of p, a McClelland-type relation (15) exists between $\langle E \rangle$, $\langle m \rangle$ and n, where the multiplyer c is determined via (16), (20) and (21).

Choosing $f(x) = 1$, the above theorem reduces to Lemma 3.

Theorem 1 shows that the existence of a McClelland-type dependency between the total π-electron energy and the topological parameters n and m is almost independent of the assumed distribution of graph eigenvalues. This also explains the success of the approximate topological formula (6) [5]. The distribution of graph eigenvalues influences, however, the value of the empirical constant a. The fact that the parameters a in (6) and b in (8) have very close numerical values indicates that the assumptions made in Lemma 3 are not very far from the reality.

Acknowledgement

One of the authors (I.G.) was an Alexander von Humboldt fellow in 1985. He thanks the Alexander von Humboldt Foundation for financial support and the staff of the Max-Planck-Institut für Strahlenchemie in Mülheim/Ruhr for assistance and hospitality.

[1] For recent work on total π-electron energy see: I. Gutman, A. V. Teodorović, and Lj. Nedeljković, Theor. Chim. Acta **65,** 23 (1984); L. Türker, Match (Mülheim) **16,** 83 (1984); Y. Jiang, A. Tang, and R. Hoffmann, Theoret. Chim. Acta **66,** 183 (1984).
[2] H. Ichikawa and Y. Ebisawa, J. Amer. Chem. Soc. **107,** 1161 (1985).
[3] For details of the topological theory of conjugated molecules and fundamentals of graph theory see: A. Graovac, I. Gutman, and N. Trinajstić, Topological Approach to the Chemistry of Conjugated Molecules, Springer-Verlag, Berlin 1977; I. Gutman and O. E. Polansky, Mathematical Aspects in Organic Chemistry, Springer-Verlag, Berlin, to appear.
[4] B. J. McClelland, J. Chem. Phys. **54,** 640 (1971).
[5] I. Gutman, Lj. Nedeljković, and A. V. Teodorović, Bull. Soc. Chim. Beograd **48,** 495 (1983); I. Gutman, A. V. Teodorović, and Ž. Bugarčić, Bull. Soc. Chim. Beograd **49,** 521 (1984).
[6] I. Gutman, Match (Mülheim) **14,** 71 (1983).