Information entropy as a measure of the correlation energy associated with the cumulant

Yu Wang,1 P. J. Knowles,2 and Jian Wang 3,*
1Department of Physics, Jiaxing University, Zhejiang 314001, China
2School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff CF10 3AT, United Kingdom
3School of Science, Huzhou University, Zhejiang 313000, China

(Received 16 May 2021; accepted 3 June 2021; published 15 June 2021)

A linear relation is found between Jaynes entropy and the correlation energy associated with the cumulant for a correlated molecular system, as demonstrated in H2 and a water molecule. The relation modifies the Collins conjecture, in which the correlation energy is defined as the total energy difference with reference to the Hartree-Fock counterpart. The linear behavior is system dependent and shifts slightly on a different basis set. This correlation energy is extensive. For a system whose wave function can be expressed as the antisymmetrized product of orthogonal groups, the correlation energy is just the sum of the individual groups. The finding offers a convenient way to construct an approximate correlation energy functional in reduced density-matrix-functional theory, which may be used in large-scale molecular dynamics simulations and potential energy surface calculations where the molecular geometry needs to be frequently updated.

DOI: 10.1103/PhysRevA.103.062808

I. INTRODUCTION

The reduced density-matrix-functional theory (RDMFT) is an extension of the density-functional theory (DFT) [1]. The fundamental variable in RDMFT is the one-particle density matrix (1-matrix for short), which has an eigenfunction expansion

\[ \gamma(1, 1') = \sum_i n_i \chi_i(1) \chi_i^*(1'), \quad (1) \]

where the eigenfunctions \( \chi_i(1) \) are called the natural orbitals. The eigenvalues \( n_i \) are their occupation numbers, which have the property \( \sum n_i = N \).

The advantage of RDMFT is that the kinetic energy is not required the construction of an approximate functional, so only the two-electron Coulomb repulsion energy need to be modeled as a functional of the 1-matrix. The two-electron repulsion energy is usually expressed as

\[ U_{ee} = \int \frac{\Gamma(1, 2)}{r_{12}} d1 d2 = \sum_{ijkl} \Gamma_{ijkl}(i j | k l), \quad (2) \]

where \( \Gamma(1, 2) \) is the diagonal elements of the two-particle density matrix (2-matrix), also called the pair density. \( (i j | k l) \) are the two-electron integrals. In the method of the configuration interaction (CI), \( \Gamma_{ijkl} \) depend on the expansion coefficients of the wave function. In RDMFT, \( \Gamma_{ijkl} \) are regarded as functionals of 1-matrix \( \gamma \) in principle. However, since the orbitals are already contained in the two-electron integrals \( (i j | k l) \), one usually assumes \( \Gamma_{ijkl} \) as only functionals of the occupation numbers \( n_i \). In practical functionals, the functionals of \( \Gamma_{ijkl} \) often take the square-root form \( \sqrt{n_i} \) [2–10], or the power-law form \( n_i^\alpha \), with \( \alpha \approx 0.656 \).

The cumulant of a 2-matrix is defined as \[ 12,13 \]

\[ \lambda(1, 2) = \Gamma(1, 2) - \frac{1}{2} [\gamma(1, 1) \gamma(2, 2) - \gamma(1, 2) \gamma(2, 1)], \quad (3) \]

where \( \gamma \) is the 1-matrix originating from the same wave function as \( \Gamma(1, 2) \). While the 2-matrix is a product-separable quantity as the wave function, it is not extensive. The cumulant has the nice property of being an extensive, additively separable entity \[ 12 \], \( \gamma(1, 1) \gamma(2, 2) \) is the pair density of independent particles, while \( \gamma(1, 2) \gamma(2, 1) \) is the exchange effect. The cumulant \( \lambda(1, 2) \) contains the effect of the Coulomb correlation.

In DFT and also in RDMFT, the effects of exchange and Coulomb correlations were often modeled together through the exchange-correlation hole \[ 3,14–17 \], which is responsible for the exchange-correlation energy. Another strategy is to treat the exchange and Coulomb correlation separately, where the two-electron repulsion energy associated with the cumulant is regarded as the correlation energy \[ 12,18 \],

\[ E_{\text{cum}} = \int \frac{\lambda(1, 2)}{r_{12}} d1 d2. \quad (4) \]

The exchange effect can be either treated in the sense of a Hartree-Fock approximation or treated in \( \gamma(1, 2) \gamma(2, 1) \), with \( \gamma \) originated from the correlated system. In quantum chemistry, there is another definition of correlation energy, which is defined as the total energy beyond the Hartree-Fock (HF) energy \[ 19 \],

\[ E_{\text{corr}} = E - E_{\text{HF}}. \quad (5) \]
The Hartree-Fock energy is the minimum for any 2-matrix of the form [20]

\[ K(1, 2) = \frac{1}{2}[\gamma(1, 1)\gamma(2, 2) - \gamma(1, 2)\gamma(2, 1)] \]  

This form is exact in the Hartree-Fock method [21], i.e., the cumulant is zero in Hartree-Fock.

Information entropy [22] is a measure of available information, complete or incomplete, given a set of possible events with probabilities \( p_1, p_2, \ldots, p_n \), where \( \sum p_i = 1 \). On the other hand, the probability distribution can be regarded as the result of a maximum-entropy principle subject to certain constraints [23]. For example, the Gaussian distribution is the result of maximum entropy subject to the condition that the standard deviation be fixed at \( \sigma \). The square of the absolute value of the wave function has a physical interpretation as a probability distribution. Information entropies based on density matrices or density have been the topic of extensive investigations for many years [24–29].

For the information entropy on the 1-matrix, Jaynes [30] has given the form

\[ S = -\sum_i n_i \ln n_i. \]  

For the Hartree-Fock approximation, since \( n_i \) are either 0 or 1, the entropy is zero. In the CI method, \( 0 < n_i < 1 \), the entropy is positive. As the correlation energy, the entropy can be regarded as a measure of correlation strength.

Collins [31] once conjectured that the correlation energy, \( E_{\text{corr}} \), with the sign reversed, be proportional to Jaynes entropy,

\[ E_{\text{corr}} = \kappa \sum_i n_i \ln n_i, \]  

where \( \kappa \) is a constant. Numerical evidence for the conjecture has since been reported in the lithium isoelectronic series and small molecules using sequences of increasingly large basis sets [32,33], while discrepancies have also been found [34,35].

The design of a RDMFT functional often faces the choice as to which 2-matrix elements to model. Most RDMFT functionals are two-index \( JK \) type, i.e., containing only \( J_{ij} = (ii|jj) \) and \( K_{ij} = (ij|ji) \) spatial two-electron integrals [36]. Such functionals can be easily regarded as incomplete in representing the 2-matrix of real molecules from the wave function, because many of the matrix elements with different \( i, j, k, l \) indices are neglected. Given the fact that our knowledge of the 2-matrix or cumulant of a real molecule is limited, what is the best guess for a RDMFT functional? The motivation of this article is to find a method which can cover all the information in the 2-matrix, maybe of a statistical nature, so we land on the concept of information entropy.

II. COMPUTATION RESULTS

Figure 1 displays the entropy \( S \) as a function of internuclear distance \( R \) (in bohrs) in \( \text{H}_2 \). The data are listed in Table I.

![FIG. 1. Jaynes entropy of \( \text{H}_2 \) as a function of internuclear distance \( R \) (in bohrs) in \( \text{H}_2 \). The data are listed in Table I.](image-url)
which also contains other essential data for our analysis. $U_{ee}$ and $U_{0ee}$ are the two-electron repulsion energies in CI and HF calculations. $Y$ is the two-electron energy calculated with the $K$ matrix defined in Eq. (6),

$$Y = \int K(1,2) r_{12} d1d2,$$

where the 1-matrix is from the correlated CI wave function.

The data were calculated with the MOLPRO code [37] using the augmented correlation-consistent polarized valence sextuple zeta basis set [aug-cc-pV6Z(s)], which gives a total energy $-1.15485$ a.u. at an equilibrium bond length $R = 1.4$ bohrs.

The first line in Table I with a bond length $R = 0$ corresponds to the He atom. The entropy for the He atom is small, which means the Hartree-Fock approximation is a good approximation.

From Eqs. (2)–(4) and (9), the correlation energy associated with the cumulant can be written as

$$E_{\text{cum}} = U_{ee} - Y.$$  

In Fig. 2, $-E_{\text{cum}}$ (with a sign change) is plotted as a function of entropy. The data fall on almost a straight line with the equation $y = 0.20493x + 0.02808$. For comparison, we also display the energy difference of $U_{0ee} - U_{ee}$ as a function of the entropy. Note that $U_{0ee}$ is closer to $U_{ee}$ than $Y$, because their difference is smaller. However, the linear behavior of $U_{0ee} - U_{ee}$ with $S$ is limited to a nuclear distance $R < 4$ bohrs.

In the original conjecture of Collins, the correlation energy is referred to as the energy difference with respect to the Hartree-Fock reference. The correlation energy associated with the cumulant has no reference to the Hartree-Fock. Both the correlation energy associated with the cumulant and the entropy come from the same CI wave function, so one may expect a closer relationship between them.

Figure 3 compares the data from two basis sets. The data from the correlation-consistent polarized valence double zeta (cc-pVDZ) basis set can be represented by the line $y = 0.19886x + 0.03687$ after linear regression. The figure indicates that the linear behavior shifts only slightly with a different basis set.

Next, we study the water molecule H$_2$O, which has three internal coordinates: two H-O bond lengths $R_1$ and $R_2$, plus the angle $\theta$ between them. Table II lists the entropies and the two-electron energies at selected geometries. The data are calculated in the complete active space self-consistent field method (CASSCF) with six electrons in 14 active orbitals. The wave function consists of 41405 configuration state functions with the cc-pVDZ basis set [38,39]. At short bond lengths, such as $R_1 = R_2 = 1.5$ bohrs, the wave function is dominated by a single determinant. As the bond length increases, the wave function evolves into a multiconfiguration character.

A linear relation between the correlation energy $-E_{\text{cum}}$ and the entropy is suggested in Fig. 4. The data fit well to a line $y = 0.220x + 0.188$. On the other hand, although $U_{ee}$ is closer

FIG. 2. The correlation energy associated with the cumulant as a function of Jaynes entropy in H$_2$. Also in display are the differences of two-electron energy between the HF and CI calculations, $U_{0ee} - U_{ee}$.

FIG. 3. The comparison of a linear relation between the correlation energy associated with the cumulant and entropy for two basis sets in H$_2$, aug-cc-pV6Z(s) and cc-pVDZ.

FIG. 4. The correlation energy associated with the cumulant is plotted with respect to Jaynes entropy in H$_2$O. For comparison, the differences of two-electron energies between HF and CI calculations, $U_{0ee} - U_{ee}$, are also in display. The data are from Table II.
to $U_{\text{loc}}$ than $Y$ in magnitude, the difference of $U_{\text{loc}} - U_{ee}$ is much scattered with respect to the entropy.

The above results suggest an approximation of the correlation energy associated with the cumulant in terms of the entropy,

$$E_{\text{cum}} = \kappa \sum \ln n_i - b,$$

(11)

where $\kappa$ and $b$ are positive constants depending on the system and the basis set, but insensitive to the geometry. Once they are calculated, one can use them as known parameters, or one may find them on the spot by two independent CI calculations to find two points on the line. The formula is thus convenient for the molecular dynamics simulation and potential energy surface calculation [40] where the geometry needs to be frequently updated.

### III. DISCUSSION AND CONCLUSION

Using the fact that the correlation energy associated with the cumulant is extensive, one may find some property of the constants $\kappa$ and $b$. For example, for a system of two $H_2$ located at a far distance, the combined system will have the same $\kappa$ but $b$ doubled. In general, for a wave function that can be approximated as the antisymmetrized product of orthogonal groups [41], the 2-matrix $\Gamma(1, 2)$ can be split into two parts: One part with the two electrons (1 and 2) in the same group, and another part with two electrons in different groups, say, the group $I$ and $J$. The contribution to $\Gamma(1, 2)$ of the latter has the structure

$$\Gamma_{IJ}(1, 2) = \gamma'((1, 1)\gamma'((2, 2) - (1, 2)\gamma'((2, 1)$$

$$- \gamma'((2, 1)\gamma'((2, 2) + \gamma'((2, 2)\gamma'((1, 1).$$

(12)

Such 1-matrix related terms have been included in the $K$ matrix in Eq. (6), and they do not contribute to the correlation energy associated with the cumulant. So the correlation energy for the orthogonal groups is just the sum of the individual groups,

$$E_{\text{cum}} = \sum_i k_i \sum n'_i \ln n'_i - \sum b_i,$$

(13)

where $k_i$, $n'_i$, and $b_i$ are the data of the individual group $I$. In terms of correlation energy, orthogonal groups are just
isolated fragments. If correlation beyond exchange comes into play when we bring chemical groups together, then new parameters $\kappa$ and $b$ need to be established.

One example of orthogonal groups is the antisymmetrized product of strongly orthogonal geminals (APSG) [12,42], and its wave function is expressed as

$$\Psi(1, \ldots, 2n) = A \prod_{m=1}^{n} \psi_m(2m-1, 2m),$$

where $A$ is the usual antisymmetrizer. While there is exchange between different geminals, correlation exists only within the individual geminal.

In the multiconfiguration self-consistent field (MCSCF) or CASSCF method, the wave function for the whole system is the product of a determinant for the closed shells and a CI form for the active space. The orbitals in the closed shells may contribute a cross term to the 2-matrix $\Gamma(1, 2)$ in the form of Eq. (12) [43], but they do not contribute to the correlation energy. Only the orbitals in the active space contribute to the correlation energy and entropy.

A close analog to our finding is the electron gas in the Thomas-Fermi approximation. From Gell-Mann and Brueckner’s work on the correlation energy of an electron gas at high density [44], Grassi et al. [45] obtained the following relation between the correlation energy per unit volume and Shannon entropy of the electron density,

$$E_{c}^{\text{eqn}} = q_0 \int n(r) \ln[n(r)]d^3r - q_1 \int n(r)d^3r,$$

where $q_0 = 0.02073$ and $q_1 = 0.00584$ in a.u. The parameters $q_0$ and $q_1$ are similar to our parameters $\kappa$ and $b$.

As Levy [18] noted, only a small piece of the exact universal variational functional of the 1-matrix is unknown, which is the correlation energy associated with the cumulant. Our result of Eq. (11) implies a simple functional of RDMFT for a correlated molecular system,

$$E = \sum_i n_i h_i + \frac{1}{2} \sum_{i,j} n_in_j\langle ij|ij\rangle + \kappa \sum_i n_i \ln n_i - b.$$ (16)

The functional contains only $JK$-type two-electron integrals as in the Hartree-Fock approximation. The contributions of other two-electron integrals are magically condensed into the entropy and the two parameters $\kappa$ and $b$. In traditional 1-matrix functionals [46,47], one often introduces complicated phase structures to resolve the phase dilemma for the 2-matrix element $\Gamma_{ij,kl}$ [48,49]. There is no need to worry about the phases here, since all the different choices or uncertainties related to the phases and the matrix elements $\Gamma_{ij,kl}$ are treated statistically in the sense of information theory based on the premise of the maximum-entropy principle. Another problem with traditional 1-matrix functionals is their nonuniqueness, which means they cannot differentiate systems with the same 1-matrix eigenvalue spectrum but a different 2-matrix, because $\Gamma_{ij,kl}$ are usually treated as the functional of the occupation numbers only [50–52]. Due to the system-dependent parameters $\kappa$ and $b$, the nonuniqueness problem can be relieved.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (Grant No. 11274109).