Natural molecular shells as open subsystems of small molecules

Alex D. Gottlieb,[1] John D. Head,[2] and Dennis Perusse[1]
[1] *Wolfgang Pauli Institute, Nordbergstrasse 15, 1090 Vienna, Austria*
[2] *Department of Chemistry, University of Hawai‘i*

(Dated: November 5, 2010)

We compare the first few natural molecular shells of several small molecules to the corresponding "RHF", "MP2", and "LDA" molecular shells. Occupation probabilities of each molecular shell are computed *ab initio*, especially the single-occupation and double-occupation probabilities, i.e., the probabilities that the molecular shell is occupied by exactly one electron or by exactly two electrons. We observe that, among corresponding molecular shells, the natural molecular shell has the least single-occupation probability and the greatest double-occupation probability.

I. INTRODUCTION

In a molecule, regarded as a many-electron system in the non-relativistic, pinned nucleus approximation, the electrons that occupy a "molecular shell" constitute an "open" many-electron system. We define a molecular shell to be a linear space of spin-orbitals that is invariant with respect to the spatial and spin symmetries of the molecule. The system of electrons that occupy a molecular shell is open in the sense that electrons can enter or leave it. If the molecular shell is d-dimensional, then the open system associated to the molecular shell can contain as few as 0 and as many as d electrons. We shall call such a many-electron system an "open subsystem" of the molecule, of rank d.

Molecular shells may be described using any system of molecular orbitals (MOs) that possess the symmetries of the molecule. Often, such MOs are obtained from a self-consistent field computation, such as a Hartree-Fock (HF) or a density functional theory (DFT) computation, and one may then consider molecular shells that are spanned by canonical HF orbitals or by Kohn-Sham (KS) orbitals. In this paper, the MOs of primary interest are the natural spin-orbitals (NSOs) introduced by Löwdin, who argued that "the introduction of natural spin orbitals leads … to a configurational expansion of most rapid convergence" [1, 2]. The NSOs are are the eigenvectors of the 1-particle reduced density matrix derived from the full wave function of the molecule. We normalize the 1-particle reduced density matrix to have trace N, and call this the 1-matrix. A diagonal matrix element $\langle\psi|\gamma|\psi\rangle$ of the 1-matrix γ is the occupation probability of the spin-orbital ψ. The occupation probabilities of the NSOs, called "natural occupation numbers", are the eigenvalues of the 1-matrix.

We call molecular shells spanned by NSOs "natural" molecular shells, and we call open subsystems associated to natural molecular shells "natural" open subsystems. From what we have seen [3], *natural* open subsystems tend to be in states of comparatively low entropy. In this article we present some empirical evidence to this effect. We perform *ab initio* computations on spin singlet ground states of several small molecules, and compare natural open subsystems to open subsystems described by, e.g., HF or KS orbitals.

Suppose that an N-electron molecule has a non-degenerate ground state (necessarily a spin singlet, so N is even). Since the ground state is non-degenerate, its 1-matrix commutes with all of the unitary operators in the 1-electron symmetry group, and therefore the eigenspaces of the 1-matrix are invariant under the group of symmetries. Thus the eigenspaces of the 1-matrix are natural molecular shells.

Consider, for example, the 6-electron molecule Li_2. The exact ground state wave function of Li_2 is unknown, but can be approximated by wave functions expressed in terms of a small number of MOs. We use the open source software

GAMESS [4], in which the MO space is spanned by a basis of contracted gaussian-type orbitals (CGTOs). The CGTO basis set pc-1 [27] gives us 18 spatial orbitals with which to describe Li_2, making the MO space 36-dimensional. Using the experimental inter-atom distance 2.6730Å [6], we compute the full configuration interaction (CI) ground state wave function

$$\Psi_{\mathrm{pc1}} \;=\; \sum_{j_1<\cdots<j_6} C_{j_1\cdots j_6}|\phi_{j_1}\,\phi_{j_2}\,\phi_{j_3}\,\phi_{j_4}\,\phi_{j_5}\,\phi_{j_6}| \tag{1}$$

with respect to an orthonormal system $\phi_1,\ldots,\phi_{36}$ of MOs. The 6-electron Slater determinants $|\phi_{j_1}\,\phi_{j_2}\,\cdots\,\phi_{j_6}|$ appearing in the CI expansion (1) are normalized, so $\sum|C_{j_1\cdots j_6}|^2=1$.

The 1-matrix derived from Ψ_{pc1} is 36×36. The eigenspace corresponding to its largest eigenvalue (0.999218) is 2-dimensional. It is spanned by α and β NSOs of σ_g symmetry. We call this eigenspace the $1\sigma_g$ natural molecular shell. The next largest eigenvalue of the 1-matrix is 0.999212. The corresponding eigenspace is spanned by α and β NSOs of σ_u symmetry; it is the $1\sigma_u$ natural molecular shell. The eigenspace corresponding to the next largest eigenvalue, 0.916467, is also 2-dimensional. It is spanned by α and β NSOs of σ_g symmetry, so we call it the $2\sigma_g$ natural molecular shell.

The occupation probability of an individual NSO is an eigenvalue of the 1-matrix. The probability that *several* NSOs are occupied *simultaneously* can be computed from the CI expansion (1) with respect to the NSOs. For example, the probability that the NSOs ϕ_1 and ϕ_2 are both occupied equals the sum of the squares of the coefficients of the configurations that contain both ϕ_1 and ϕ_2. In this way we can determine that the probability is about 8.3% that no electrons would be found in the $2\sigma_g$ molecular shell of Ψ_{pc1}. This information cannot be obtained from the 1-matrix, as it concerns a 2-particle correlation.

In the present article, we report occupation probabilities of the first few natural molecular shells of the molecules Be, LiH, BH, BeH_2, and Li_2. These all have non-degenerate spin singlet ground states and are small enough for full CI calculations to be convenient. The results depend considerably upon the basis set of CGTOs employed for the computation, so we have tried more than one basis set with each molecule.

We are going to compare each natural molecular shell to molecular shells that are described by other kinds of spin-orbitals. For each molecule and basis set considered, we also perform a restricted Hartree-Fock (RHF) computation, a restricted second-order Møller-Plesset (MP2) computation, and a restricted DFT computation with a local density approximation (LDA) density functional. We thereby obtain three additional systems of orthonormal MOs: the canonical RHF orbitals, the LDA Kohn-Sham orbitals, and the eigenvectors of the MP2 "response density" matrix. These orbitals have the symmetries of the molecule in all of the cases we have investigated, and so we are able to identify "RHF", "MP2", and "LDA" molecular shells spanned by these orbitals. The RHF and LDA molecular shells are eigenspaces of the 1-electron Hartree-Fock and LDA Kohn-Sham Hamiltonians, respectively, just as the natural molecular shells are eigenspaces of the 1-matrix [7].

We use the full CI wave function to compute the occupation probabilities of the RHF, MP2, and LDA molecular shells. Thus, the occupation numbers of the RHF molecular shells are not simply integers, as they are in the Hartree-Fock approximation, since we are using the full CI wave function rather than the HF Slater determinant to compute them. Similarly, the occupations of the MP2 molecular shells are not those of the MP2 level of theory, but those given by the best available representation of the exact many-electron wave function. Our purpose here is not to contrast full CI with approximate levels of theory, but to show that *natural* molecular shells are extraordinary. We just use the RHF, MP2, and LDA orbitals to describe molecular shells that we can meaningfully contrast with the natural molecular shells.

Returning to the example of Li_2, let us now compare its natural, RHF, MP2, and LDA $2\sigma_g$ molecular shells. From the full CI expansion with the basis set pc-1, we compute the probabilities p_2, p_1, and p_0 that the $2\sigma_g$ molecular shell is occupied by $2, 1$, or 0 electrons, respectively. We repeat the same computations using the basis set tzv [5]. Table 1 below reports the computed occupation probabilities and the average occupation $p_1 + 2p_2$ of each molecular shell. "Entropy" is the von Neumann entropy of the state of the open subsystem, as given by formula (14) in Section II D, and "size" refers to the number of spatial orbitals in the basis set.

Table 1: occupations of the $2\sigma_g$ molecular shells of the Li_2 ground state

Basis Set	Size		Entropy	Avg. Occ.	p_2	p_1	p_0
pc-1	18	Natural	0.4211	1.83293	0.91618	0.00058	0.08324
		RHF	0.4380	1.83139	0.91459	0.00222	0.08320
		MP2	0.4227	1.83284	0.91606	0.00073	0.08321
		LDA	0.4252	1.83280	0.91591	0.00099	0.08311
tzv	26	Natural	0.4426	1.82195	0.91054	0.00087	0.08859
		RHF	0.4567	1.82062	0.90918	0.00226	0.08856
		MP2	0.4431	1.82190	0.91049	0.00092	0.08859
		LDA	0.4495	1.82151	0.90997	0.00157	0.08846

Occupation probabilities differ considerably between the two basis sets. However, for each of the basis sets, observe that the natural molecular shell has the greatest "double-occupation" probability p_2, the least "single-occupation" probability p_1, and the least von Neumann entropy.

This article presents the results of similar computations on five small molecules. For each molecule, we compare the occupation probabilities of the first few corresponding natural, RHF, MP2, and LDA molecular shells. We judge that molecular shells "correspond" to each other if (i) they were computed using the same basis set, (ii) they have the same symmetry index, e.g., $2\sigma_g$, (iii) they have comparable average occupations, and (iv) the correspondence is unambiguous. In the present study, it was easy to identify sets of corresponding heavily occupied molecular shells of all four kinds (NSO, RHF, MP2, and LDA). However, the lightly occupied molecular shells often did not permit themselves to be placed into correspondence with one another (see the discussion around Table 7 in Section IV B 2).

We observe that, in most sets of four corresponding molecular shells, the natural molecular shell has the greatest double-occupation probability and the least single-occupation probability. We believe that these observations manifest a general virtue of natural molecular shells, namely, that the open subsystems described by natural molecular shells tend to be in states of high internal order. In Table 1, the internal order of the states is reflected in the distribution of their occupation probabilities p_0, p_1, and p_2. One may quantify the degree of internal *disorder* of a molecular shell's state by its von Neumann entropy, defined below in Section II C, and indeed it appears that natural open subsystems tend to be in states of minimal von Neumann entropy. In this article we focus on occupation probabilities p_0, p_1, and p_2 more than von Neumann entropies, because the von Neumann entropy is just a summary statistic, whereas the occupation probabilities provide a more detailed view of what is going on in the molecular shells.

The rest of this paper is organized into four sections. In Section II we review the theoretical framework for many-electron systems, and define open subsytems and their occupation probabilities. In Section III we muse on the advantages of using natural molecular shells for describing electronic structure, and cite a couple of mathematical propositions. In Section IV, we describe our computational methods, report the data, and try to explain what we have observed using ideas from Section III. Section V is a brief conclusion.

II. FORMALISM AND TERMINOLOGY

The Hilbert space $\mathcal{H}$ for a single electron is isomorphic to $L^2(\mathbb{R}^3 \times \{\alpha, \beta\})$. That is, vectors in $\mathcal{H}$ may be represented by wave functions $\psi(x, \sigma)$, where x ranges over space ($\mathbb{R}^3$) and σ ranges over the two spin states (α and β). Unit vectors in $\mathcal{H}$ are called "spin-orbitals" or simply "orbitals". A pure state of an N-electron molecule is represented by a wave function in the N-electron Hilbert space, which is isomorphic to $\mathcal{H}^{\wedge N}$, the N-fold exterior power of $\mathcal{H}$. Vectors in $\mathcal{H}^{\wedge N}$ may be represented by wave functions $\Psi(x_1, \ldots, x_N)$ that are antisymmetric in the variables $x_i = (\mathbf{r}_i, \sigma_i)$.

The "1-matrix" γ derived from Ψ is the operator on the 1-electron Hilbert space $\mathcal{H}$ with integral kernel

$$\gamma(x, y) \;=\; N \int\!\!\int \cdots \int \Psi(x, x_2, \ldots, x_N)\Psi^*(y, x_2, \ldots, x_N)\; dx_2 \cdots dx_N \tag{2}$$

(where the integrals include summation over the spin indices). Löwdin called γ the "density matrix of order 1" [1]. The eigenvectors of the 1-matrix γ are called "natural spin-orbitals" (NSOs) and its eigenvalues are called "natural occupation numbers." The probability that a (normalized) spin-orbital ψ is occupied equals $\langle \psi | \gamma | \psi \rangle$.

A. CI expansions and occupation probabilities

Given any orthonormal basis $\{\phi_1, \phi_2, \ldots\}$ of $\mathcal{H}$, the set $\Big\{ |\phi_{j_1}\, \phi_{j_2} \cdots \phi_{j_N}| : j_1 < j_2 < \cdots < j_N \Big\}$ of "Slater determinants" in the spin-orbitals ϕ_i is an orthonormal basis of $\mathcal{H}^{\wedge N}$. Thus, every wave function Ψ in $\mathcal{H}^{\wedge N}$ has a unique "configuration interaction" (CI) expansion

$$\Psi \;=\; \sum_{j_1 < \cdots < j_N} C_{j_1 \cdots j_N} |\phi_{j_1}\, \phi_{j_2} \cdots \phi_{j_N}| \tag{3}$$

with respect to any given ordered orthonormal basis $(\phi_1, \phi_2, \ldots)$ of $\mathcal{H}$.

The joint occupation probability of several (orthonormal) spin-orbitals can be computed most conveniently from a CI expansion with respect to any orthonormal basis that contains those spin-orbitals. For example, the probability that ϕ_i and ϕ_j are both occupied is the sum of the squares of the coefficients of the configurations in (3) that contain both ϕ_i and ϕ_j. Although we *compute* occupation probabilities this way, they can be *defined* in a way that is independent of the basis used in the CI expansion. We do this in the following paragraphs.

B. Open subsystems

Molecular shells determine open many-electron systems. The molecular shell itself is a finite-dimensional subspace $\mathcal{K}$ of the 1-electron Hilbert space $\mathcal{H}$. It is open in the sense that electrons can move in and out of it. The "open subsystem" consisting of the electrons that happen to be in $\mathcal{K}$ will be said to have "rank" $d = \dim(\mathcal{K})$. The Hilbert space for this open subsystem is the fermion Fock space over $\mathcal{K}$, which has dimension 2^d. The fermion Fock space over $\mathcal{K}$ is

$$\mathcal{F}_{\mathcal{K}} \;=\; \mathbb{C} \oplus \mathcal{K} \oplus \mathcal{K}^{\wedge 2} \oplus \mathcal{K}^{\wedge 3} \oplus \cdots \oplus \mathcal{K}^{\wedge d}\ . \tag{4}$$

The Fock space $\mathcal{F}_{\mathcal{K}}$ has $d+1$ components, one for each number of electrons that could possibly be present in the open subsystem. The first component, i.e., $\mathbb{C}$, is a 1-dimensional Hilbert space that accomodates the contingency that no electrons are present in the system. The last component, i.e., $\mathcal{K}^{\wedge d}$, is also 1-dimensional, for there is only one way for d fermions to occupy d modes.

Statistical “states” of the open subsystem with Hilbert space $\mathcal{F}_{\mathcal{K}}$ are represented by density operators D on $\mathcal{F}_{\mathcal{K}}$. When the open subsystem is in the state given by D, the probability that it is occupied by exactly n electrons is

$$p_n \;=\; \mathrm{Tr}(DP_n)\ , \tag{5}$$

where P_n denotes the orthogonal projector onto the n-electron component of $\mathcal{F}_{\mathcal{K}}$. The average number of electrons in the system is $\sum np_n$. We call this the “average occupation” of the open subsystem.

The density operators we consider all have the block-diagonal form

$$D \;=\; D_0 \oplus D_1 \oplus D_2 \oplus \cdots \oplus D_d \tag{6}$$

with respect to the decomposition (4). This condition means that the many-electron state is a statistical mixture of states of definite electron number. For such states $p_n = \mathrm{Tr}(D_n)$.

The electronic state of the entire molecule determines the state of each its open subsystems, as follows. Let $\mathcal{K}$ be a closed subspace of the 1-electron Hilbert space $\mathcal{H}$. The Fock space over $\mathcal{H}$ is isomorphic to the tensor product of the Fock spaces over $\mathcal{K}$ and $\mathcal{K}^{\perp}$, the ortho-complement of $\mathcal{K}$ [9, 10], that is,

$$\mathcal{F}_{\mathcal{H}} \cong \mathcal{F}_{\mathcal{K}} \otimes \mathcal{F}_{\mathcal{K}^{\perp}}\ . \tag{7}$$

Formula (7) means that the entire system (the electrons in the molecule) can be broken into two open subsytems: one consisting of all the electrons in $\mathcal{K}$ and the other consisting of all the electrons *not* in $\mathcal{K}$. If Ψ is an N-electron wave function, representing the state of the *entire* many-electron system, then the state of the open *subsystem* is represented by the density operator

$$D = \mathrm{Tr}_{\mathcal{F}_{\mathcal{K}^{\perp}}} |\Psi\rangle\langle\Psi| \tag{8}$$

on $\mathcal{F}_{\mathcal{K}}$, where $\mathrm{Tr}_{\mathcal{F}_{\mathcal{K}^{\perp}}}$ denotes the partial trace over $\mathcal{F}_{\mathcal{K}^{\perp}}$. It can be shown (but we omit the proof) that D has the block-diagonal form (6). It can also be shown that

$$\sum np_n \;=\; \sum_{i=1}^{d} \langle\psi_i|\gamma|\psi_i\rangle \tag{9}$$

is the average occupation of the open subsystem, where γ is the 1-matrix (2) of Ψ and $\{\psi_1, \ldots, \psi_d\}$ is any orthonormal basis of $\mathcal{K}$.

The partial trace in (8) is the unique operator D on $\mathcal{F}_{\mathcal{K}}$ such that $\mathrm{Tr}(DX) = \langle\Psi|\widehat{X}\Psi\rangle$ for all bounded operators X on $\mathcal{F}_{\mathcal{K}}$, where $\widehat{X}$ denotes the operator on $\mathcal{F}_{\mathcal{H}}$ defined via the isomorphism (7) by

$$\widehat{X}(\phi \otimes \psi) \;=\; (X\phi) \otimes \psi \quad \text{for all} \quad \phi \in \mathcal{F}_{\mathcal{K}}, \psi \in \mathcal{F}_{\mathcal{K}^{\perp}}.$$

Since the trace $\mathrm{Tr}(DX)$ is independent of the orthonormal basis used to compute it, the *partial* trace (8) is also basis-independent. Thus, formulae (8) and (5) provide a basis-independent definition of the occupation probabilities.

C. The von Neumann entropy

The “von Neumann entropy” of a density operator D is [11]

$$S(D) \;=\; -\mathrm{Tr}(D \log_2 D) \;=\; -\sum \lambda_i \log_2 \lambda_i\ , \tag{10}$$

where $\lambda_1, \lambda_2, \ldots$ are the eigenvalues of D. Since D is positive semi-definite and has trace 1, its eigenvalues are nonnegative and sum to 1. The von Neumann entropy of D is thus the "Shannon entropy" of the discrete probability distribution defined by the eigenvalues of D.

In quantum information theory, the von Neumann entropy of a reduced density matrix such as (8) is the definitive measure of the "entanglement" between two parties who share a pure quantum state [12]. According to formula (7), the electrons in a molecule can be regarded as a "bipartite" quantum system, composed of the open subsystem $\mathcal{F}_{\mathcal{K}}$ and a complementary open subsystem $\mathcal{F}_{\mathcal{K}^\perp}$. This way of thinking about entanglement in many-fermion systems was set forth by Zanardi [13] in the context of fermionic lattices, and was later adopted by Huang and Kais [24] to study electron correlation in molecules. However, Wiseman and Vaccaro [25] point out that much of this so-called "entanglement of modes" is not available as a resource for quantum information processing, since the particle number super-selection rule (cf., formula (6)) precludes many of the operations that would be required to access it. Therefore, we do not refer to the von Neumann entropy $S(D)$ as the "entanglement" of $\mathcal{F}_{\mathcal{K}}$ and $\mathcal{F}_{\mathcal{K}^\perp}$. We are only using the von Neumann entropy as a kind of summary statistic to express the degree of disorder in the state of the open subsystem $\mathcal{F}_{\mathcal{K}}$.

D. Open subsystems of rank 2

Most of the open sytems investigated in the present paper consist of (the electrons in) 2-dimensional molecular shells. In such cases, $\mathcal{K} \cong \mathbb{C}^2$ is spanned by α and β spin-orbitals with the same spatial wave function.

Let $\mathcal{K}$ denote a 2-dimensional subspace of the 1-electron Hilbert space, with an orthornormal basis $\{\psi_1, \psi_2\}$. Then

$$\mathcal{F}_{\mathcal{K}} \cong \mathbb{C} \oplus \mathcal{K} \oplus \mathbb{C} \tag{11}$$

is only 4-dimensional. States of the open subsystem $\mathcal{F}_{\mathcal{K}}$ are represented by density matrices D on $\mathcal{F}_{\mathcal{K}}$ that have block-diagonal form

$$D \;=\; D_0 \oplus D_1 \oplus D_2 \;=\; [p_0] \oplus \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \oplus [p_2] \tag{12}$$

relative to the decomposition (11). p_0 is the probability of "double vacancy" and p_2 is the probability of "double occupation." The matrix D_1 contains information about the event where there is *exactly one* electron in $\mathcal{K}$. For example, the probability that there would be found only one electron, and that it would be in the state represented by the unit vector $\vec{u} = c_1\psi_1 + c_2\psi_2$, is

$$\langle \vec{u} | D_1 \vec{u} \rangle \;=\; \begin{bmatrix} \bar{c}_1 & \bar{c}_2 \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \;. \tag{13}$$

The trace of D_1 is the probability p_1 that exactly one electron would be found in span$\{\psi_1, \psi_2\}$. Since p_1 is the probability that there is exactly one electron in the open subsystem and p_2 is the probability that there are two electrons in it, the average occupation of the open subsystem is $p_1 + 2p_2$.

The 2-dimensional molecular shells considered in this paper are spanned by α and β spin-orbitals with the same spatial wave function, and their states are derived as in formula (8) from spin singlet states of the entire molecule. Because of the spin symmetry, the two eigenvalues of D_1 are both equal to $\frac{1}{2}p_1$. Therefore the eigenvalues of the density matrix (12) are $p_0, \frac{1}{2}p_1, \frac{1}{2}p_1$, and p_2; and its von Neumann entropy equals

$$-p_0 \log_2(p_0) - p_1 \log_2(p_1/2) - p_2 \log_2(p_2) \;. \tag{14}$$

III. NATURAL OPEN SUBSYSTEMS

What is the best way to describe the electronic structure of an N-electron molecule using only d spin-orbitals? Let us imagine that one may use any orbitals, but only d of them, say $d \approx N$. One may try to approximate the N-electron wave function Ψ by a wave function that only uses the spin-orbitals one has selected. This is what is done in the Hartree-Fock approach, for example, where $d = N$ and the wave function in d spin-orbitals which satisfies the "best energy criterion" is sought [16]. This is a good way to proceed if energy is the only observable of interest. However, for more general "data compression" purposes, to get the most informative description of Ψ, taking fullest advantage of the d available spin-orbitals, it might be better to use the density matrix (8) on the Fock space over the span of the selected spin-orbitals. For one thing, the information contained in this density matrix is exact, that is, it exactly describes all correlations of all orders between the selected orbitals. This may not always be the *best* way to use d spin-orbitals to describe Ψ, but it is an option worth considering.

Suppose, then, that one wishes to compress the information in an N-electron wave function Ψ and just store a reduced density matrix (8) of the form (6). Let us imagine again that d is fixed and relatively small, but that one may select *any* d orbitals. Which d orbitals should one select? In other words, which open subsystem of rank d should one choose? A sensible option would to select the open subsystem whose average occupation is as great as possible.

The open subsystem of rank d containing the greatest average number of electrons is the natural open subsystem associated to the d most heavily occupied NSOs. This is a consequence of the following general theorem [14, 15]:

Proposition 1 *Let* Γ *be a Hermitian operator satisfying* $\mathrm{Tr}(\Gamma^2) < \infty$*, and let* $\lambda_1 \geq \lambda_2 \geq \cdots$ *denote its eigenvalues. Then*

$$\sum_{i=1}^{d} \lambda_i \;\geq\; \sum_{i=1}^{d} \langle \psi_i | \Gamma | \psi_i \rangle \tag{15}$$

for any d *orthonormal vectors* $\psi_1, \ldots, \psi_d$*.*

When $\Gamma = \gamma$, the 1-matrix derived from Ψ, the right-hand side of inequality (15) is the average occupation of $\mathrm{span}\{\psi_1, \ldots, \psi_d\}$ and the left-hand side is the average occupation of the span of the d most heavily occupied NSOs (cf., formula (9)). Moreover, Proposition 1 also implies that the average occupation of the most heavily occupied natural molecular shell of any symmetry species is at least as large as the average occupation of any other molecular shell of the same symmetry. To prove this, note that γ is block-diagonal with respect to the decomposition of the 1-electron space into symmetry components, and apply Proposition 1 with Γ equal to the *restriction* of γ to one of its symmetry components.

The mathematically demonstrable advantages of the NSOs mainly proceed [17] from the fact that they are the eigenvectors of the 1-matrix, and therefore [18] minimize concave functions of the 1-orbital occupation probabilities. The results we just derived from Proposition 1 are of this type, for they concern average occupation number, which is a 1-particle observable.

The present study, however, has to do with *correlations* between particles, such as the probability p_1 that one spin-orbital of a given pair is occupied while the other is vacant. This is a 2-particle statistic, which cannot be determined by the 1-matrix alone. In the present study of the first several molecular shells of small atoms, we observe that p_1 tends to be smaller in natural molecular shells. To explain this would require more information than is contained in the 1-matrix. The following proposition about the natural CI expansion of "geminals" [19, 20] is relevant in this regard:

Proposition 2 *Let Φ be a unit vector in a 2-fermion Hilbert space $\mathcal{H} \wedge \mathcal{H}$. Then there exist orthonormal vectors $\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots$ in $\mathcal{H}$ such that*

$$\Phi \;=\; c_1 \; |\alpha_1 \, \beta_1| \;+\; c_2 \; |\alpha_2 \, \beta_2| \;+\; \cdots \tag{16}$$

The vectors α_i and β_i are natural orbitals for Φ with natural occupation numbers $|c_i|^2$.

For example, if the natural CI expansion (16) for the ground state of H_2 is ordered so that $|c_1| \geq |c_2| \geq |c_3| \geq \cdots$, then α_1 and β_1 would be the $1\sigma_g$ spin-orbitals and α_2 and β_2 would be the $1\sigma_u$ spin-orbitals. Of course, the conventional CI expansion using HF spin-orbitals also begins with $1\sigma_g$ and $1\sigma_u$ orbitals; the important point here is that the *natural* $1\sigma_g$ and $1\sigma_u$ spin-orbitals do not appear in any other configurations in (16).

Proposition 2 tells us everything we need to know about correlations between the NSOs in the case of 2-electron wave functions. Consider a 2-electron spin singlet wave function Φ with natural CI expansion (16). It is impossible for a natural molecular shell to be occupied by exactly one electron, since the form (16) of the wave function ensures that β_j is occupied if α_j is, and vice versa. For this reason we have not included data for 2-electron molecules like H_2 in this article: there is *never* just one electron in a natural molecular shell of the H_2 ground state and therefore p_1 is exactly equal to 0 in every natural molecular shell!

If versions of Proposition 2 were available for 4-electron and 6-electron wave functions, they might help explain why p_1 tends to be smallest in the natural molecular shells. Unfortunately, only a couple of very closely related results are available [21].

IV. DATA

A. Computational Methods

We use the open source software GAMESS [4] to approximate the ground states of the molecules Be, LiH, BH, BeH_2, and Li_2. GAMESS computes with basis sets of CGTOs. We have used the basis sets tzv [5], cc-pVDZ and cc-pVTZ [8], and pc-1 and pc-2 [26, 27].

For the diatomic molecules LiH, BH, and Li_2, we took the experimental inter-nuclear distances 1.5957Å, 1.2324Å, and 2.6730Å , respectively; and for the linear molecule H-Be-H, we took the experimental Be-H bond length 1.3400Å [6].

For each molecule and basis set, we performed RHF, second-order Møller-Plesset, and Kohn-Sham LDA computations to obtain the RHF, MP2, and LDA MOs. The Kohn-Sham LDA calculations used the Slater [28] exchange and VWN-5 [29] correlation density functionals, and the default numerical grids in GAMESS. The MP2 MOs are the eigenvectors of the MP2 "response density matrix" [30], which GAMESS computes for us and calls "MP2 natural orbitals." Note that the response density matrix is not derived from any wavefunction as in formula (2). In fact, it need not be a density matrix at all — in some cases it is not even positive semi-definite [31]. Nonetheless, in the cases considered in this paper, the MP2 MOs are usually extremely close to the NSOs.

We performed full CI computations using the RHF, MP2, and LDA MOs as reference orbitals. These CI computations provide the natural orbitals as a by-product. We then performed another CI computation to obtain the "natural" CI expansion.

We inspected the MOs in order to identify their symmetries and classify the molecular shells. We then computed the occupation probabilities of the molecular shells from the CI expansions. Since GAMESS reports CI expansion coefficients to 7 decimal places, the occupation probabilities we computed are probably only accurate to 5 or 6 decimal places. The occupation probabilities reported in the following tables have been rounded to five decimal places.

Most of the molecular shells we studied are 2-dimensional, and one may use formula (14) to compute the von Neumann entropies directly from the occupation probabilities. However, to compute the von Neumann entropies for the 6-dimenstional 2p shells of Be and the 4-dimensional 1π shells of BH and Li_2, it was actually necessary to compute the partial trace (8) to obtain the reduced density matrix (6). Entropies are only reported for these molecular shells, rounded to 4 decimal places.

B. Tables

The format of the the following tables is the same as for Table 1 in the Introduction, except that entropies are only reported in Tables 4, 11, and 17. Exclamation points and asterisks in the tables designate exceptional cases, to be discussed in Section IV C.

1. Be

Table 2: occupation probabilities of 1s subshells of the Be ground state

Basis Set	Size		Avg. Occ.	p_2	p_1	p_0
tzv	14	Natural	1.99837	0.99907	0.00023	0.00070
		RHF	1.99834	0.99904	0.00027	0.00070
		MP2	1.99834	0.99905	0.00025	0.00071
		LDA	1.99778	0.99846	0.00085	0.00068
pc-1	9	Natural	1.99919	0.99948	0.00024	0.00028
		RHF	1.99915	0.99944	0.00027	0.00029
		MP2	1.99916	0.99944	0.00028	0.00028
		LDA	1.99882	0.99912	0.00058	0.00030
pc-2	16	Natural	1.99937	0.99957	0.00023	0.00020
		RHF	1.99934	0.99954	0.00025	0.00020
		MP2	1.99933	0.99953	0.00028	0.00020
		LDA	1.99886	0.99906	0.00073	0.00021
cc-pVDZ	15	Natural	1.99978	0.99986	0.00007	0.00007
		RHF	1.99975	0.99982	0.00010	0.00007
		MP2	1.99974	0.99982	0.00011	0.00007
		LDA	1.99918	0.99925	0.00068	0.00007
cc-pVTZ	35	Natural	1.99824	0.99896	0.00031	0.00073
		RHF	1.99821	0.99893	0.00035	0.00072
		MP2	1.99821	0.99894	0.00033	0.00073
		LDA	1.99765	0.99836	0.00093	0.00071

Table 3: occupation probabilities of 2s subshells of the Be ground state

Basis Set	Size		Avg. Occ.	p_2	p_1	p_0
tzv	14	Natural	1.80997	0.90487	0.00023	0.09490
		RHF	1.80874	0.90356	0.00161	0.09483
		MP2	1.80952	0.90438	0.00076	0.09486
		LDA	1.80960	0.90437	0.00086	0.09477
pc-1	9	Natural	1.84638	0.92308	0.00023	0.07669
		RHF	1.84491	0.92149	0.00192	0.07659
		MP2	1.84590	0.92253	0.00084	0.07663
		LDA	1.84565	0.92217	0.00131	0.07652
pc-2	16	Natural	1.82300	0.91140	0.00020	0.08840
		RHF	1.82141	0.90971	0.00199	0.08830
		MP2	1.82257	0.91090	0.00076	0.08833
		LDA	1.82253	0.91078	0.00097	0.08825
cc-pVDZ	15	Natural	1.81668	0.90831	0.00007	0.09162
		RHF	1.81572	0.90728	0.00117	0.09156
		MP2	1.81641	0.90799	0.00044	0.09158
		LDA	1.81620	0.90769	0.00083	0.09148
cc-pVTZ	35	Natural	1.81894	0.90932	0.00029	0.09039
		RHF	1.81746	0.90776	0.00194	0.09030
		MP2	1.81862	0.90898	0.00067	0.09035
		LDA	1.81839	0.90865	0.00110	0.09026

Table 4 reports the occupation probabilities of 2p molecular subshells of the Be ground state. Because there are only 4 electrons in Be, the occupation probabilities p_5 and p_6 of a 6-dimensional 2p subshell of Be must equal 0. The occupation probabilities p_3 and p_4 need not equal 0 exactly, but they all happen to equal 0.00000 when rounded to five decimal places, and therefore they are not tabulated. The von Neumann entropy is tabulated, however, because it is not simply a function of p_0, p_1, and p_2 as in formula (14).

Table 4: occupation probabilities of 2p subshells of the Be ground state

Basis Set	Size		Entropy	Avg. Occ.	p_2	p_1	p_0
tzv	14	Natural	0.4491	0.18642	0.09314	0.00013	0.90672
		RHF	0.5950	0.14389	0.05598	0.03192	0.91210
		MP2	0.4528	0.18608	0.09280	0.00047	0.90673
		LDA	0.5130	0.17651	0.08365	0.00920	0.90714
pc-1	9	Natural	0.3880	0.15005	0.07490	0.00025	0.92485
		RHF	0.3883	0.15001	0.07486	0.00028	0.92486
		MP2	0.3889	0.14993	0.07479	0.00035	0.92486
		LDA	0.3919	0.14971	0.07453	0.00064	0.92483
pc-2	16	Natural	0.4282	0.17337	0.08658	0.00020	0.91322
		RHF	0.4398	0.17206	0.08531	0.00144	0.91325
		MP2	0.4288	0.17331	0.08652	0.00025	0.91323
		LDA	0.4284	0.17336	0.08657	0.00021	0.91322
cc-pVDZ	15	Natural	0.4369	0.17973	0.08984	0.00005	0.91011
		RHF	0.5693	0.14727	0.06058	0.02611	0.91331
		MP2	0.4386	0.17959	0.08970	0.00019	0.91011
		LDA	0.4852	0.17299	0.08330	0.00639	0.91031
cc-pVTZ	35	Natural	0.4340	0.17676	0.08828	0.00020	0.91152
		RHF	0.5755	0.12153	0.04214	0.03724	0.92062
		MP2	0.4378	0.17638	0.08791	0.00056	0.91153
		LDA	0.5026	0.16556	0.07758	0.01039	0.91203

2. **LiH**

Table 5: occupation probabilities of 1σ molecular shells of the LiH ground state

Basis Set	Size			Ave. Occ.	p_2	p_1	p_0
tzv	16		Natural	1.99477	0.99719	0.00039	0.00242
			RHF	1.99475	0.99717	0.00041	0.00242
			MP2	1.99477	0.99719	0.00039	0.00242
			LDA	1.99425	0.99665	0.00095	0.00240
pc-1	14		Natural	1.99852	0.99914	0.00024	0.00062
		!	RHF	1.99851	0.99913	0.00024	0.00063
			MP2	1.99851	0.99913	0.00026	0.00062
			LDA	1.99803	0.99866	0.00071	0.00063
pc-2	31		Natural	1.99751	0.99857	0.00037	0.00106
			RHF	1.99748	0.99855	0.00038	0.00107
			MP2	1.99750	0.99856	0.00038	0.00106
			LDA	1.99702	0.99808	0.00085	0.00107
cc-pVDZ	20		Natural	1.99956	0.99974	0.00009	0.00017
			RHF	1.99955	0.99973	0.00010	0.00017
			MP2	1.99956	0.99974	0.00009	0.00017
			LDA	1.99873	0.99889	0.00095	0.00016
cc-pVTZ	50		Natural	1.99629	0.99794	0.00040	0.00165
			RHF	1.99627	0.99791	0.00044	0.00165
		*!	MP2	1.99629	0.99794	0.00040	0.00165
			LDA	1.99550	0.99713	0.00125	0.00163

Table 6: occupation probabilities of 2σ molecular shells of the LiH ground state

Basis Set	Size		Ave. Occ.	p_2	p_1	p_0
tzv	16	Natural	1.95359	0.97661	0.00037	0.02302
		RHF	1.95017	0.97301	0.00416	0.02284
		MP2	1.95285	0.97583	0.00119	0.02298
		LDA	1.95308	0.97605	0.00099	0.02297
pc-1	14	Natural	1.95779	0.97878	0.00023	0.02099
		RHF	1.95570	0.97664	0.00241	0.02094
		MP2	1.95739	0.97836	0.00067	0.02097
		LDA	1.95627	0.97708	0.00211	0.02081
pc-2	31	Natural	1.94674	0.97319	0.00036	0.02645
		RHF	1.94611	0.97255	0.00102	0.02643
		MP2	1.94663	0.97307	0.00049	0.02644
		LDA	1.94385	0.96999	0.00386	0.02614
cc-pVDZ	20	Natural	1.94712	0.97352	0.00008	0.02640
		RHF	1.94614	0.97250	0.00113	0.02637
		MP2	1.94696	0.97333	0.00028	0.02639
		LDA	1.94493	0.97108	0.00276	0.02616
cc-pVTZ	50	Natural	1.94166	0.97063	0.00039	0.02897
		RHF	1.94124	0.97020	0.00084	0.02896
		MP2	1.94160	0.97056	0.00049	0.02895
		LDA	1.93875	0.96744	0.00386	0.02869

We will not compare further shells, e.g., the 3σ and 1π shells, because it becomes unclear how to do so. In Table 7a we list the average occupations of the natural molecular shells for the basis cc-pcVDZ, and in Table 7b we list the average occupations of the RHF molecular shells (these are listed in order of increasing HF energies, which are not reported here). Although one may easily identify the 3σ natural molecular shell, it is not clear whether any σ RHF molecular shell can be said to "correspond" to it. The average occupation of the 3σ natural molecular shell is approximately 0.031, but no single RHF molecular shell has a comparable average occupation. Instead, there are two RHF molecular shells of σ symmetry, both of which have average occupation of about 0.009. The average occupation of the 1π natural molecular shell is more than twice as large as that of the most heavily occupied RHF molecular shell of π symmetry, and therefore we doubt it would make sense to compare them.

Table 7a: natural molecular shells for LiH/cc-pcVDZ

	1σ	2σ	3σ	1π	4σ	5σ	6σ	2π
Ave. Occ.	1.99956	1.94712	0.03145	0.01468	0.00570	0.00036	0.00032	0.00061

	7σ	8σ	3π	4π	9σ	5π	10σ
Ave. Occ.	0.00013	0.00002	0.00003	0.00002	0.00000	0.00000	0.00000

Table 7b: RHF molecular shells for LiH/cc-pcVDZ

	1σ	2σ	σ	π	σ	σ	π	σ
Ave. Occ.	1.99955	1.94614	0.00501	0.00172	0.008772	0.00019	0.00287	0.00886

	π	π	σ	σ	σ	π	σ
Ave. Occ.	0.00747	0.00002	0.00270	0.00106	0.01139	0.00326	0.00100

3. **BH**

Table 8: occupation probabilities of 1σ molecular shells of the BH ground state

Basis Set	Size		Ave. Occ.	p_2	p_1	p_0
tzv	17	Natural	1.99904	0.99944	0.00016	0.00040
		RHF	1.99903	0.99943	0.00017	0.00040
		MP2	1.99903	0.99943	0.00017	0.00040
		LDA	1.99866	0.99905	0.00056	0.00039
pc-1	20	Natural	1.99963	0.99975	0.00012	0.00013
		RHF	1.99960	0.99973	0.00015	0.00013
		MP2	1.99959	0.99972	0.00015	0.00013
		LDA	1.99923	0.99935	0.00052	0.00012
cc-pVDZ	20	Natural	1.99981	0.99986	0.00009	0.00005
		RHF	1.99979	0.99984	0.00011	0.00005
		MP2	1.99979	0.99984	0.00011	0.00005
		LDA	1.99944	0.99949	0.00046	0.00004

Table 9: occupation probabilities of 2σ molecular shells of the BH ground state

Basis Set	Size		Ave. Occ.	p_2	p_1	p_0
tzv	17	Natural	1.97033	0.98194	0.00644	0.01162
		RHF	1.95023	0.96287	0.02450	0.01264
		*! MP2	1.97016	0.98205	0.00606	0.01189
		LDA	1.94781	0.96087	0.02608	0.01306
pc-1	20	Natural	1.96352	0.97811	0.00730	0.01459
		RHF	1.94824	0.96277	0.02271	0.01453
		*! MP2	1.96317	0.97804	0.00708	0.01488
		LDA	1.94527	0.96005	0.02516	0.01479
cc-pVDZ	20	Natural	1.95828	0.97479	0.00871	0.01651
		RHF	1.94387	0.96015	0.02357	0.01628
		! MP2	1.95807	0.97472	0.00863	0.01665
		LDA	1.94019	0.95672	0.02675	0.01653

Table 10: occupations of valence 3σ molecular shells of the BH ground state

Basis Set	Size		Ave. Occ.	p_2	p_1	p_0
tzv	17	Natural	1.86717	0.93051	0.00615	0.06334
		RHF	1.88582	0.93048	0.02486	0.04466
		! MP2	1.86697	0.93042	0.00614	0.06345
		! LDA	1.88804	0.93164	0.02475	0.04361
pc-1	20	Natural	1.86828	0.93069	0.00689	0.06242
		RHF	1.88202	0.92924	0.02353	0.04723
		MP2	1.86832	0.93064	0.00705	0.06231
		LDA	1.88463	0.93061	0.02342	0.04597
cc-pVDZ	20	Natural	1.85746	0.92462	0.00822	0.06716
		RHF	1.87059	0.92327	0.02404	0.05269
		MP2	1.85740	0.92446	0.00847	0.06707
		LDA	1.87330	0.92427	0.02476	0.05097

Table 11 reports the occupation probabilities of the 1π molecular shells of BH. These are 4-dimensional molecular shells, which may contain as many as 4 electrons. The average occupation is $p_1 + 2p_2 + 3p_3 + 4p_4$.

Table 11: occupations of 1π nonbonding molecular shells of the BH ground state

Basis Set	Size		Entropy	Avg. Occ.	p_4	p_3	p_2	p_1	p_0
tzv	17	Natural	0.3344	0.11881	0.00008	0.00000	0.05888	0.00074	0.94031
		RHF	0.4068	0.08413	0.00002	0.00003	0.02986	0.02423	0.94586
		MP2	0.3368	0.11850	0.00008	0.00000	0.05857	0.00104	0.94030
		LDA	0.3598	0.11443	0.00007	0.00001	0.05469	0.00475	0.94049
pc-1	20	Natural	0.3647	0.11970	0.00008	0.00004	0.05766	0.00394	0.93828
		RHF	0.4158	0.09907	0.00005	0.00005	0.03988	0.01892	0.94109
		MP2	0.3677	0.11918	0.00007	0.00005	0.05716	0.00444	0.93828
		LDA	0.3893	0.11418	0.00009	0.00003	0.05240	0.00892	0.93856
cc-pVDZ	20	Natural	0.4030	0.12906	0.00006	0.00008	0.06103	0.00652	0.93231
		RHF	0.4563	0.10706	0.00007	0.00007	0.04172	0.02313	0.93501
		* MP2	0.4024	0.12869	0.00006	0.00008	0.06085	0.00653	0.93248
		LDA	0.4436	0.11994	0.00011	0.00004	0.05184	0.01569	0.93232

4. **BeH_2**

Table 12: occupations of valence $1\sigma_g$ molecular shells of the BeH_2 ground state

Basis Set	Size		Ave. Occ.	p_2	p_1	p_0
tzv	20	Natural	1.99831	0.99903	0.00024	0.00072
		RHF	1.99830	0.99902	0.00025	0.00072
		MP2	1.99831	0.99903	0.00025	0.00072
		LDA	1.99776	0.99847	0.00082	0.00071
pc-1	19	Natural	1.99916	0.99946	0.00025	0.00029
		RHF	1.99915	0.99945	0.00025	0.00030
		MP2	1.99916	0.99946	0.00025	0.00029
		LDA	1.99881	0.99911	0.00059	0.00030
cc-pVDZ	25	Natural	1.99968	0.99978	0.00013	0.00009
		RHF	1.99967	0.99976	0.00015	0.00009
	*	MP2	1.99968	0.99978	0.00013	0.00009
		LDA	1.99911	0.99920	0.00071	0.00009

Table 13: occupations of valence $2\sigma_g$ molecular shells of the BeH_2 ground state

Basis Set	Size		Ave. Occ.	p_2	p_1	p_0
tzv	20	Natural	1.96658	0.97751	0.01156	0.01093
		RHF	1.96598	0.97688	0.01222	0.01090
		MP2	1.96648	0.97741	0.01167	0.01092
		LDA	1.96460	0.97543	0.01375	0.01083
pc-1	19	Natural	1.96543	0.97636	0.01271	0.01093
		RHF	1.96523	0.97614	0.01294	0.01092
		MP2	1.96539	0.97632	0.01275	0.01093
		LDA	1.96206	0.97283	0.01639	0.01077
cc-pVDZ	25	Natural	1.95601	0.96844	0.01914	0.01243
		RHF	1.95576	0.96817	0.01943	0.01240
		MP2	1.95599	0.96841	0.01916	0.01243
		LDA	1.95246	0.96476	0.02293	0.01230

Table 14: occupations of valence $1\sigma_u$ molecular shells of the BeH_2 ground state

Basis Set	Size		Ave. Occ.	p_2	p_1	p_0
tzv	20	Natural	1.96898	0.97882	0.01133	0.00985
		RHF	1.96886	0.97871	0.01144	0.00985
		MP2	1.96896	0.97881	0.01135	0.00985
		LDA	1.96858	0.97843	0.01172	0.00985
pc-1	19	Natural	1.96709	0.97727	0.01255	0.01018
		RHF	1.96688	0.97707	0.01275	0.01018
		MP2	1.96707	0.97725	0.01256	0.01018
		LDA	1.96665	0.97683	0.01298	0.01019
cc-pVDZ	25	Natural	1.95678	0.96881	0.01917	0.01203
		RHF	1.95669	0.96873	0.01924	0.01203
		MP2	1.95677	0.96880	0.01917	0.01203
		LDA	1.95644	0.96848	0.01949	0.01204

5. $\mathbf{Li_2}$

Table 15: occupations of $1\sigma_g$ molecular shells of the Li_2 ground state

Basis Set	Size		Avg. Occ.	p_2	p_1	p_0
pc-1	18	Natural	1.99844	0.99875	0.00094	0.00031
		RHF	1.99842	0.99873	0.00095	0.00032
		MP2	1.99842	0.99873	0.00097	0.00031
		LDA	1.99804	0.99837	0.00131	0.00032
tzv	26	Natural	1.99471	0.99592	0.00287	0.00121
		RHF	1.99469	0.99590	0.00289	0.00121
		! MP2	1.99470	0.99592	0.00287	0.00121
		LDA	1.99426	0.99546	0.00334	0.00120

Table 16: occupations of $1\sigma_u$ molecular shells of the Li_2 ground state

Basis Set	Size		Avg. Occ.	p_2	p_1	p_0
pc-1	18	Natural	1.99842	0.99874	0.00094	0.00032
		RHF	1.99842	0.99874	0.00095	0.00032
		! MP2	1.99842	0.99874	0.00094	0.00032
		LDA	1.99772	0.99803	0.00165	0.00032
tzv	26	Natural	1.99471	0.99592	0.00287	0.00121
		RHF	1.99468	0.99590	0.00289	0.00121
		! MP2	1.99471	0.99592	0.00287	0.00121
		LDA	1.99401	0.99521	0.00359	0.00120

The occupation probabilities of the $2\sigma_g$ molecular shells have already been displayed in Table 1 in the Introduction.

The next table reports the occupation probabilities of 1π molecular shells. The occupation probabilities p_3 and p_4 all equal 0.00000 when rounded to five decimal places and are not tabulated.

Table 17: occupations of 1π molecular shells of the Li_2 ground state

Basis Set	Size		Entropy	Avg. Occ.	p_2	p_1	p_0
pc-1	18	Natural	0.2747	0.09289	0.04631	0.00027	0.95342
		* RHF	0.2747	0.09288	0.04630	0.00027	0.95342
		MP2	0.2749	0.09283	0.04626	0.00031	0.95343
		LDA	0.2770	0.09270	0.04609	0.00052	0.95339
tzv	26	Natural	0.3088	0.10908	0.05441	0.00025	0.94534
		RHF	0.3144	0.10831	0.05369	0.00093	0.94538
		MP2	0.3090	0.10902	0.05437	0.00028	0.94535
		LDA	0.3096	0.10902	0.05434	0.00034	0.94533

Table 18: occupations of $2\sigma_u$ molecular shells of the Li_2 ground state

Basis Set	Size		Avg. Occ.	p_2	p_1	p_0
pc-1	18	Natural	0.04037	0.02015	0.00006	0.97979
		RHF	0.03068	0.01192	0.00683	0.98125
		MP2	0.04035	0.02013	0.00008	0.97979
		LDA	0.03455	0.01473	0.00510	0.98017
tzv	26	Natural	0.03681	0.01816	0.00049	0.98135
		RHF	0.02531	0.00900	0.00731	0.98369
		MP2	0.03678	0.01808	0.00063	0.98130
		LDA	0.03012	0.01237	0.00538	0.98225

Table 19: occupations of $3\sigma_g$ molecular shells of the Li_2 ground state

Basis Set	Size		Avg. Occ.	p_2	p_1	p_0
pc-1	18	Natural	0.03341	0.01662	0.00016	0.98321
		RHF	0.03337	0.01632	0.00074	0.98295
		MP2	0.03334	0.01656	0.00023	0.98322
		LDA	0.03112	0.01427	0.00258	0.98315
tzv	26	Natural	0.03182	0.01580	0.00022	0.98398
		RHF	0.02993	0.01386	0.00220	0.98394
		MP2	0.03172	0.01567	0.00037	0.98396
		LDA	0.02794	0.01220	0.00354	0.98426

C. Discussion

The tables show that p_1 of a natural molecular shell is smaller than p_1 of the corresponding molecular shells, and its p_2 is greater, with few exceptions. The exceptional cases are indicated in the tables with exclamation points. We first review these exceptional cases.

There is only one exception among the RHF shells, marked by an exclamation point in Table 5, where p_1 of the RHF/pc-1 1σ molecular shell in LiH is actually smaller than p_1 of the corresponding natural molecular shell. (In the table, these single-occupation probabilities have both been rounded to the fifth decimal place, and are reported to be 0.00024, but to six decimal places they are 0.000243 and 0.000239.) There is only one exception among the LDA shells: in Table 10, p_2 of the LDA/tzv 3σ molecular shell of BH is larger than p_2 of the corresponding natural molecular shell. The other eight exceptions involve the MP2 molecular shells. All but one of the exceptional cases occur when the non-natural molecular shell resembles the natural molecular shell very closely. To assess how closely the shells resemble one another we compute the overlap between the α spin-orbital of the non-natural molecular shell and the corresponding NSO (the exceptional shells are 2-dimensional and each one contains only one α spin-orbital). The overlaps are are all greater than 0.998, except for the LDA case in Table 9, where $|\langle\alpha_{\text{LDA}}|\alpha_{\text{NSO}}\rangle| \approx 0.891$.

The observation that p_2 is largest in the natural molecular shells is not surprising. Except for the 3σ shell of BH, the molecular shells we have studied are the most heavily occupied or the second most heavily occupied of their symmetry species. Proposition 1 implies that the average occupation of the most heavily occupied natural molecular shell of any symmetry species is at least as large as the average occupation of any other molecular shell of the same symmetry. The proposition further suggests – though it does not strictly imply – that the average occupation of the *second* most heavily occupied natural molecular shell of a given symmetry species will be larger than that of the second most heavily occupied RHF, MP2, or LDA molecular shell of the same symmetry. Since the average occupations of the most heavily occupied natural molecular shells are largest, one may expect their double-occupation probabilities to be largest as well.

The question is, why is p_1 *smallest* in the natural molecular shells? Proposition 2 suggests a possible answer. According to Proposition 2, natural molecular shells of 2-electron wave functions have $p_1 = 0$. In the 4-electron molecules Be and LiH, two of the four electrons are tied up in the core shells, leaving just two electrons to occupy the other natural molecular shells. The single-occupation probability p_1 of the natural $2\sigma_g$ shell isn't *exactly* 0 because the $1\sigma_g$ shell isn't quite fully occupied, but it remains relatively small. This argument might also apply to Li_2, where four of the six electrons are tied up in the core orbitals, but it is much less compelling for the 6-electron molecules BH and BeH_2.

We also find that the von Neumann entropy (10) is usually smallest in the natural open subsystems. Exceptional cases in our data, where the von Neumann entropy of a non-natural open subsystem is *smaller* than that of the corresponding natural open subsystem, are indicated in the tables by asterisks. We have only reported the values of the entropy for the 4-dimensional 1π shells of BH and Li_2 (Tables 11 and 17) and the 6-dimensional 2p shells of Be (Table 4). The rest of the molecular shells are 2-dimensional and the von Neumann entropy be computed directly from p_2, p_1, and p_0 using formula (14).

V. CONCLUSION

We performed *ab initio* computations on spin singlet ground states of 4-electron and 6-electron molecules and compared molecular shells described by NSOs to those described by three other kinds of MOs: canonical RHF orbitals, restricted LDA Kohn-Sham orbitals, and eigenvectors of the MP2 response density matrix.

Among corresponding molecular shells, the single-occupation probability of the natural molecular shell is almost always the smallest, and its double-occupation probability the largest. The natural open systems are almost always in states of lower von Neumann entropy than the open systems associated to the corresponding RHF, LDA, and MP2 molecular shells.

We speculate that we have been observing instances of a more general phenomenon, i.e., that even in larger molecules, natural open systems will generally tend to be in states of comparatively low entropy.

Acknowledgments A.G. is supported by the Vienna Science and Technology Fund project "Correlation in Quantum Systems" and D.P. was supported by a stipend from the Wolfgang Pauli Institute. A.G. and D.P. thank the University of Hawai'i Mathematics Department for use of its facilities. We thank Robert Hammerling and Frank Verstraete for helpful comments.

---

[1] Löwdin, P.-O. Quantum theory of many-particle systems I: physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configuration interaction. *Phys Rev* 1955, 97, 1474 - 1489.

[2] Löwdin. P.-O.; Shull, H. Natural orbitals in the quantum theory of two-electron systems. *Phys Rev* 1956, 101, 1730 - 1739.

[3] In the present study, and also when we sample random pure states of, e.g., 5 fermions in 10 modes. [A.D. Gottlieb and R.M. Weishäupl, unpublished work.]

[4] Schmidt, M.W.; Baldridge, K.K.; Boatz, J.A.; Elbert, S.T.; Gordon, M.S.; Jensen, J.H.; Koseki, S.; Matsunaga, N.; Nguyen, K.A.; Su, S.J.; Windus, T.L.; Dupuis, M.; Montgomery, J.A.; General Atomic and Molecular Electronic Structure System. *J Comput Chem* 1993, 14, 1347 - 1363.

[5] Dunning,T.H. Gaussian basis functions for use in molecular calculations I: contractions of (9s,5p) atomic basis sets for the first-row atoms. *J Chem Phys* 1970, 53, 2823 - 2833.

[6] NIST Standard Reference Database 101. Computational Chemistry Comparison and Benchmark DataBase. National Institute of Standards and Technology, 2009.

[7] These eigenspaces will generally be invariant under the point group and spin symmetries, i.e., they will be molecular shells. The eigenspaces of the LDA Kohn-Sham Hamiltonian must be molecular shells if the final electron density is symmetric. The eigenspaces of the RHF Hamiltonian are necessarily molecular shells if the molecule has a *unique* RHF ground state.

[8] Dunning, T.H. Gaussian basis sets for use in correlated molecular calculations I: the atoms boron through neon and hydrogen. *J Chem Phys* 1989, 90, 1007 - 1023.

[9] Parthasarathy, K.R. An Introduction to Quantum Stochastic Calculus; Birkhäuser Verlag: Basel, 1992; Prop. 19.7

[10] Alicki, R.; Fannes, M. Quantum Dynamical Systems; Oxford University Press: New York, 2001; Cap. 6

[11] Alicki, R.; Fannes, M.*op. cit.*, Chapter 9.

[12] Donald, M.J.; Horodecki, M.; Rudolph, O. The uniqueness theorem for entanglement measures. *J Math Phys* 2002, 43, 4252–4272.

[13] Zanardi, P. Quantum entanglement in fermionic lattices. *Phys Rev A* 2002, 65, 042101.

[14] Proof omitted. For the case where Γ acts on a finite dimensional space, see Prop. 4.1.5 of Ref. [15].

[15] Marcus, M.M.; Minc, H. A Survey of Matrix Theory and Matrix Inequalities; Dover Publications: New York,1992.

[16] Kutzelnigg, W.; Smith, V.H. Jr. On different criteria for the best independent-particle model approximation. *J Chem Phys* 1964, 41, 896 - 897.

[17] For example, the argument given in the "Note added in proof" in P.-O. Löwdin, *op. cit.*, is of this type.

[18] See Section II.4 of M.M. Marcus and H. Minc, *op. cit.*

[19] Coleman, A.J. Structure of fermion density matrices. *Rev Mod Phys* 1963, 35, 668 - 687; cf., Theorems 3.1 and 4.1

[20] Coleman, A.J.; Yukalov, V.I. Reduced Density Matrices: Coulson's Challenge (Vol. 72 of *Lecture Notes in Chemistry*); Springer-Verlag: Berlin 2001.

[21] By "particle-hole duality", Proposition 2 also gives the natural CI expansion of a wave function of N fermions (electrons) in a space of $N+2$ modes (spin-orbitals). The structure of the natural CI expansion of a wave functions for $N = 3$ fermions in a space of 6 modes is also known [22, 23].

[22] Borland, R.E.; Dennis, K. The conditions on the one-matrix for three-body fermion wavefunctions with one-rank equal to six. *J Phys B: At Mol Phys* 1972, 5, 7 - 15.

[23] Ruskai, M.B. Connecting N-representability toWeyls problem: the one-particle density matrix for $N = 3$ and $R = 6$. *J Phys A: Math Theor* 2007, 40, F961 - F967.

[24] Huang, Z.; Kais, S. Entanglement as measure of electron-electron correlation in quantum chemistry calculations. *Chem Phys Lett* 2005, 413, 1- 5.

[25] Wiseman. H.M.; Vaccaro, J. A. Entanglement of indistinguishable particles shared between two parties. *Phys Rev Lett* 2003, 91, 097902.

[26] Jensen, F. Polarization consistent basis sets: Principles. *J Chem Phys* 2001, 115, 9113 - 9125.

[27] Jensen, F. Polarization consistent basis sets VI: the elements He, Li, Be, B, Ne, Na, Mg, Al, Ar. *J Phys Chem A* 2007, 111, 11198 - 11204.

[28] Slater, J.C. A simplification of the Hartree-Fock method. *Phys Rev* 1951, 81, 385 - 390.

[29] Vosko, S.H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can J Phys* 1980, 58, 1200 - 1211.

[30] Trucks, G.W.; Salter, E.A.; Sosa, C.; Bartlett, R.J. Theory and implementation of the MBPT density matrix: an application to one-electron properties. *Chem Phys Lett* 1988, 147, 359 - 366.

[31] Gordon, M.S.; Schmidt, M.W.; Chaban, G.M.; Glaesemann, K.R.; Stevens, W.J.; Gonzalez, C. A natural orbital diagnostic for multiconfigurational character in correlated wave functions. *J Chem Phys* 1999, 110, 4199 - 4207.