Kohn-Sham Density Functional Theory with Complex, Spin-Restricted Orbitals: Accessing a New Class of Densities without the Symmetry Dilemma

Joonho Lee,† Luke W. Bertels,‡ and Martin Head-Gordon†
College of Chemistry, University of California, Berkeley, California 94720, USA.

We show that using complex, spin-restricted orbitals (cR) in Kohn-Sham density functional theory (KS-DFT) allows one to access a new class of densities that is not accessible by either spin-restricted (RKS) or spin-unrestricted (UKS) orbitals. We further show that the real part of a cRKS density is non-identical when the imaginary part of the density matrix is not zero. Using cRKS orbitals shows significant improvements in the triplet-singlet gaps of the TS12 benchmark set for the SPW92, PBE, BLYP, and TPSS functionals. Moreover, it was shown that RKS and UKS yield qualitatively wrong charge densities and spin densities, respectively, leading to worse energetics. We demonstrate that the modern functionals SCAN, MN15-L, and B97M-V show surprisingly no improvement even with a qualitatively more accurate density from cRKS orbitals. To this end, our work not only paves the way to escape the symmetry dilemma whenever there exists a cRKS solution, but also suggests a new route to design better approximate XC functionals.

Introduction

In Kohn-Sham density functional theory (KS-DFT), it is clear that for systems with antiferromagnetically coupled open-shell electrons, spin-unrestricted KS-DFT (UKS) behaves better than spin-restricted KS-DFT (RKS) at the expense of spin-polarization.1–3 For finite-sized systems, this spin-polarization is a manifestation of the lack of strong (or static) correlation in a single-determinant wavefunction.4,5 This is known as a symmetry dilemma in KS-DFT, so called because exact solutions for finite systems should not break spin-symmetry yet the unphysical, spin-symmetry-broken approximate solutions often provide lower energies than symmetric approximate solutions.

As an attempt to understand this symmetry dilemma, Perdew and co-workers proposed the use of “on-top” pair density,5,6 Π(r), which is defined as

$$\Pi(r) = \rho_\alpha(r)\rho_\beta(r).$$ (1)

Instead of working with spin density variables, ρ_α(r) and ρ_β(r), the fundamental variables are now the charge density, ρ(r) = ρ_α(r) + ρ_β(r), and the magnetization density, m(r), which is a function of ρ(r) and Π(r). One can accomplish the one-to-one mapping between m(r) and Π(r) following

$$m(r) = \rho(r) \left(1 - \frac{4\Pi(r)}{\rho(r)^2}\right)^{1/2}$$ (2)

Evidently, inserting Eq. (1) into Eq. (2) yields a more familiar definition of the magnetization density, m(r) = ρ_α(r) − ρ_β(r). With ρ(r) and m(r), one can back out ρ_α(r) and ρ_β(r) and these can be used to compute the exchange-correlation (XC) potential, v_{xc}(r). This framework allows one to treat m(r) as an auxiliary variable that is closely related to the on-top pair density. Therefore, it helps to explain spin-polarization in KS-DFT in terms of the on-top pair density of a wavefunction. While this is certainly useful in understanding this symmetry dilemma, it is somewhat unsatisfying that one has to interpret the magnetization density merely as an auxiliary variable. Moreover, it does not remedy problems arising from evaluating v_{xc}(r) with spin densities.

In this Letter, we describe a way to access a new class of charge and spin densities that is not accessible by spin-polarization. This is achieved by breaking time-reversal symmetry and complex symmetry of the KS-DFT determinant.7–10 We will refer to this symmetry breaking as “complex-polarization”. For semi-local functionals, we show that this complex-polarization does not pose any dilemma in obtaining charge and spin densities. We illustrate how this is achieved and demonstrate the numerical performance of popular semi-local functionals on chemical systems where complex-polarization is essential to obtain correct charge and spin densities.

Theory

We start from the KS-DFT Lagrangian,11

$$\mathcal{L}[\mathbf{P}] = E_{\text{KS-DFT}}[\mathbf{P}] + \mu \left(\text{tr}(\mathbf{P}) - \frac{N}{2}\right) + \text{tr}(\mathbf{A}(\mathbf{P}^2 - \mathbf{P})) + \text{tr}(\mathbf{B}(\mathbf{P}^4 - \mathbf{P}))$$ (3)

where N is the number of electrons, P is a one-particle density matrix (1PDM), μ, A, and B are Lagrange multipliers for constraining the trace, idempotency, and hermiticity of P, respectively. The idempotency of P guarantees the non-interacting nature of the KS-DFT problem. For simplicity, we also assume that the computational basis is orthogonal.

These three constraints are naturally imposed by the definition of P,

$$\mathbf{P} = \mathbf{C}_{\text{occ}} \mathbf{C}_{\text{occ}}^\dagger$$ (4)

where C_{occ} is the occupied molecular orbital (MO) coefficient matrix. However, when minimizing Eq. (3) with respect to P directly as is done in linear-scaling DFT approaches, it is necessary to consider these constraints explicitly or impose them through density matrix purification techniques.12 We write P in terms of its real and imaginary components,

$$\mathbf{P} = \mathbf{X} + i\mathbf{Y}$$ (5)
We can then rewrite these constraints as follows: the trace condition leads to
\[
\text{tr}(X) = \frac{N}{2} \tag{6}
\]
\[
\text{tr}(Y) = 0, \tag{7}
\]
the idempotency becomes
\[
X = X^2 - Y^2 \tag{8}
\]
\[
Y = XY + YX, \tag{9}
\]
and the hermiticity yields
\[
X = X^T \tag{10}
\]
\[
Y = -Y^T. \tag{11}
\]
Based on Eq. (11), we conclude that Y is antisymmetric and therefore it automatically satisfies Eq. (7). This is an important observation since \(\rho(r)\) is computed from P following
\[
\rho(r) = 2 \sum_{\mu\nu} \eta_\mu(r)\eta_\nu(r)P_{\mu\nu} \tag{12}
\]
where \(\eta_\mu(r)\) is a real-valued computational basis function. This real-valuedness is sufficient for studying bound states of molecules and indeed most basis sets for finite-sized systems are real-valued. Since \(\eta_\mu(r)\eta_\nu(r)\) is a symmetric tensor, an antisymmetric tensor, Y, does not contribute to \(\rho(r)\). In other words, it is equivalent to write
\[
\rho(r) = 2 \sum_{\mu\nu} \eta_\mu(r)\eta_\nu(r)X_{\mu\nu}.
\]

The key insight here is that by having non-zero Y, the idempotency constraint on X can be relaxed to Eq. (8) and Eq. (9). In other words, X does not have to satisfy \(X = X^2\) as long as Y is non-negligible. Such a density matrix P with \(\text{tr}(Y^T Y) \neq 0\) is referred to as a “fundamentally complex” density matrix. It is fundamentally complex in the sense that no unitary rotation of a real-valued MO coefficient matrix in the complex plane can represent a fundamentally complex density matrix. Consequently, the use of fundamentally complex density matrices results in a broader class of densities which may come from a non-idempotent density matrix X. It is not obvious whether lifting this idempotency constraint on X would always yield a lower energy solution. In fact, the energy lowering turns out to be quite rare to observe and systems with a high point group symmetry tend to yield two half-occupied orbitals. Depending on \(\theta\), it is possible to obtain a non-idempotent X. While the hidden MC form in the cRKS determinant is limited, its real PDM (for \(E_T\)) and its \(\rho(r)\) (for \(E_J\) and \(E_V\)) can all be consistently evaluated using existing XC functionals. Thus cRKS can be viewed as a limited case of MC-DFT without the formal challenges and practical double-counting problems normally associated with that field.

Results We evaluate the numerical performance of RKS, UKS, and cRKS on the recently developed TS12 benchmark set. This data set contains the experimental singlet-triplet gaps of 12 atoms and molecules: C,
NF, NH, NO\textsuperscript{−}, O\textsubscript{2}, O, PF, PH, S\textsubscript{2}, S, Si, and SO. The ground states of these molecules are triplets. The lowest singlet states for each system are singlet biradicals and exhibit a spin-restricted Hartree-Fock (RHF) to complex RHF (cRHF) instability. This instability is driven by the underlying point group symmetry which gives rise to the degeneracy between the highest occupied MO (HOMO) and the lowest unoccupied MO (LUMO).

The correct lowest singlet states of these systems should singly-occupy both HOMO and LUMO and thereby obtain charge and spin densities that obey the underlying point group symmetry. Even with the exact XC functional, RKS orbitals are qualitatively wrong in these cases as they doubly-occupy the HOMO, breaking the point group symmetry of the system and leading to a broken-symmetry charge density. Since the exact XC functional yields not only exact energy but also exact charge density, accessing a different class of density matrix other than those from RKS orbitals is necessary. A similar phenomenon was first pointed out by Pople, Gill, and Handy in the context of spin-polarization in open-shell systems.\textsuperscript{19} Indeed, UKS can achieve these single occupations by breaking spin-symmetry. However, we will show that spin-polarization is not the way to access the right charge and spin densities in the systems in TS12.

We investigate a local density approximation (LDA) functional, SPW92,\textsuperscript{20–23} two generalized gradient approximation (GGA) functionals, BLYP\textsuperscript{24,25} and PBE,\textsuperscript{26} and four meta-GGA (mGGA) functionals, TPSS, SCAN,\textsuperscript{28} MN15-L,\textsuperscript{29} and B97M-V.\textsuperscript{30} We used aug-cc-pVQZ basis set,\textsuperscript{31,32} 99 points for XC radial quadrature and 590 points for XC angular quadrature. We only report KS-DFT solutions found to be locally stable based on stability analysis.\textsuperscript{33,34} The triplet ground states are computed with UKS ($M_S = 1$) and we focus on the symmetry dilemma of the singlet states. All calculations were performed with a development version of Q-Chem.\textsuperscript{35}

| SPW92 | PBE | BLYP | TPSS |
|---|---|---|---|
| RKS | | | |
| MSD | 10.85 | 13.09 | 9.94 | 13.85 |
| RMSD | 11.46 | 13.53 | 10.34 | 14.42 |
| UKS | | | |
| MSD | -13.70 | -16.57 | -17.61 | -17.26 |
| RMSD | 14.46 | 17.16 | 18.64 | 18.19 |
| cRKS | | | |
| MSD | -1.23 | 3.00 | 0.90 | 7.94 |
| RMSD | 2.19 | 3.41 | 1.91 | 8.63 |

Table I. The deviation (kcal/mol) with respect to experimental values in $\Delta E_{L,S}(= E_S - E_T)$ of the SPW92, PBE, BLYP, and TPSS functionals obtained with restricted (RKS), unrestricted (UKS), and complex, restricted (cRKS) orbitals. RMSD stands for root-mean-square-deviation and MSD stands for mean-signed-deviation.

In Table I, we present the root-mean-square-deviation (RMSD) and mean-signed deviation (MSD) of the relatively old XC functionals SPW92, BLYP, PBE, and TPSS. Using RKS orbitals clearly overestimates the triplet-singlet gap as restricted orbitals cannot describe the open-shell nature of the lowest singlet state of these systems. These RKS energies for singlet states are too high so that the gap is overestimated, as evidenced by the positive MSDs. Using UKS orbitals has the opposite problem. Namely, the gaps are all underestimated. This is due to the undesired mixing of the singlet and triplet states (i.e., spin-contamination). Since the triplet state is lower in energy, the resulting UKS $M_S = 0$ state energy is too low. The negative MSDs are due to this spin-contamination. The use of cRKS orbitals improves both the MSD and RMSD by more than a factor of 5 in the case of SPW92, PBE and BLYP. The improvement in TPSS is less impressive although it is still a factor of 2 improvement.

This significant improvement is simply due to obtaining $\rho$ from a non-idempotent $\mathbf{X}$ \textit{without spin-polarization}. Indeed, in all systems in TS12, the resulting $\mathbf{X}$ has exactly two eigenvalues of 0.5 when obtained from these four XC functionals. These two eigenvalues represent the single-occupancy of HOMO and LUMO which is qualitatively correct for these systems. To understand what drives the energy lowering from RKS to cRKS, we decompose the energy difference between RKS and cRKS into individual energy contributions. We present this energy decomposition for $\mathbf{X}$ in Table II. The qualitative behavior discussed here holds for other systems as well. All four XC functionals show an increase in the kinetic energy as well as the XC energy. The increase in the XC energy is driven by the increase in the exchange energy. The correlation energy shows only a small change. The RKS to cRKS instability is driven by the classical Coulomb energy. In particular, a significant energy lowering in the electron-nuclear attraction energy is the driving force of this instability in all systems in the TS12 set.

The results for the more modern mGGA functionals SCAN, MN15-L, and B97M-V, are presented in Table III. The RKS and UKS results are qualitatively similar to what was observed in the above four functionals. Namely, the overestimation in gaps was observed with RKS while the underestimation in gaps was observed with UKS. SCAN with RKS orbitals shows a particularly poor behavior compared to all the other functionals examined in this work. SCAN becomes comparable to other functionals when used with UKS orbitals. The cRKS results with

$\begin{array}{|c|c|c|c|c|c|c|c|}
\hline
& \Delta E_T & \Delta E_V & \Delta E_j & \Delta E_{Coul} & \Delta E_X & \Delta E_C & \Delta E_{XC} \\
\hline
\text{SPW92} & 12.89 & -45.70 & 11.65 & -34.05 & 7.88 & 0.20 & 8.09 \\
\text{PBE} & 10.69 & -39.57 & 7.06 & -32.51 & 12.13 & -1.38 & 10.76 \\
\text{BLYP} & 9.96 & -36.43 & 4.09 & -32.33 & 13.05 & -0.38 & 12.67 \\
\text{TPSS} & 5.95 & -26.95 & -2.28 & -29.22 & 18.45 & -1.65 & 16.80 \\
\hline
\end{array}$

Table II. The energy differences (kcal/mol) between cRKS and RKS solutions of $\mathbf{X}$ (i.e., $\Delta E = \Delta E_{cRKS} - \Delta E_{RKS}$) for each component in Eq. (13). $\Delta E_{Coul} = \Delta E_j + \Delta E_V$, $\Delta E_X$, and $\Delta E_C$ are the energy differences in classical Coulomb energy, exchange energy, and correlation energy, respectively.

$\Delta E_T$
Table III. Same as Table I except we have the SCAN, MN15-L, and B97M-V functionals presented here. Note that MN15-L and B97M-V have no stable cRKS solutions for a half of the data points (i.e., C, NF, NH, NO, O2, and O).

|        | SCAN | MN15-L | B97M-V |
|--------|------|--------|--------|
| RKS    | 19.50| 10.64  | 11.82  |
| MSD    | 19.82| 11.16  | 12.22  |
| RMSD   | -16.24| -10.79| -13.69 |
| UKS    | 17.74| 11.97  | 14.71  |
| cRKS   | 15.55| 10.50  | 10.34  |
| MSD    | 16.39| 11.00  | 11.15  |

Table III. Same as Table I except we have the SCAN, MN15-L, and B97M-V functionals presented here. Note that MN15-L and B97M-V have no stable cRKS solutions for a half of the data points (i.e., C, NF, NH, NO, O2, and O).

these three modern functionals are rather surprising due to a qualitatively different behavior from that of the four functionals discussed before. The use of cRKS orbitals for these functionals does not improve the quantitative energetics. Furthermore, MN15-L and B97M-V exhibit no stable cRKS solutions for C, NF, NH, NO, O2, and O. Additionally, MN15-L shows occupation numbers of (0.61, 0.39) and (0.78, 0.22) for PF and PH, respectively. Non-idempotent X is essential to describe the open-shell nature of singlet ground states in these systems. Therefore, with this improved density matrix and ρ therefrom, approximate XC functionals should perform better. The poor performance of these modern mGGA functionals with cRKS orbitals suggests that data such as the TS12 set may be useful for XC functional development.

Lastly, we present density plots to compare the qualitative differences between densities from RKS and cRKS orbitals. As mentioned throughout this Letter, RKS orbitals break the point group spatial symmetry of systems in the TS12 set whereas cRKS orbitals preserve the symmetry.

In Figure 1, we present the real-space density of O2 computed with BLYP. The discussion also applies to other systems in the TS12 set and real-space densities are qualitatively similar across different XC functionals. O2 is cylindrically symmetric around its interatomic axis. Therefore, qualitatively correct density should exhibit this symmetry. This is found to be true with the cRKS density in Figure 1 (a). However, in Figure 1 (b), the RKS density is elongated along y-axis and overall not cylindrically symmetric. This is due to the spatial symmetry breaking caused by doubly occupying HOMO which mixes two low-lying singlet states, $^1\Delta_g$ and $^1\Sigma_g^+$. The UKS α-spin density also breaks the cylindrical symmetry to a smaller extent than does RKS and is elongated along x-axis. The UKS charge density (i.e., the sum of α and β density) is only x-y symmetric and breaks the cylindrical symmetry. The x-y symmetry is because the UKS β-spin density is rotated 90° from the α density (i.e., elongated along y axis).

The qualitatively improved UKS charge density (compared to the RKS one) is attractive though it is still qualitatively wrong. The on-top pair density interpretation provides a good way to understand the spin-symmetry breaking in this case, but there is no quantitative fit. The resulting energetics from the UKS spin-densities are far from chemical accuracy (1 kcal/mol) as it is clear in Tables I and II. In particular, due to the spatial symmetry breaking in each spin density and charge density, the KS-DFT XC energy evaluated with these spin densities are inadequate to reliably estimate the triplet-singlet gaps. In contrast to this, we observe a qualitatively correct charge density (and also a spin density which is a half of the charge density) with the cRKS determinant. At the same time, the resulting cRKS triplet-singlet gaps are improved as it is shown in Table I.

Conclusions In this Letter, we showed that it is possible to access a different class of (both charge and spin) densities that is not possible to obtain within either RKS or UKS. This is achieved by using cRKS to obtain the density. Although the cRKS determinant breaks time-reversal symmetry and complex symmetry, the resulting densities do not exhibit this symmetry breaking. Furthermore, based on the triplet-singlet test set (TS12), we showed that the cRKS charge densities follow the
point group symmetry while RKS does not. This allows cRKS to improve the quantitative accuracy of some XC functionals by a factor of 5. We also showed that the UKS charge density is qualitatively incorrect and the resulting energies are far from chemical accuracy due to spin-contamination. Lastly, we note that modern mGGA functionals (SCAN, MN15-L, and B97M-V) do not show any significant improvements even when correct densities from cRKS are used. Even with the exact XC functional, one needs a cRKS determinant to obtain qualitatively correct charge and spin densities for the systems considered here. This Letter suggests that these modern functionals might lack some aspects of the exact XC functional. We hope that our study provides a new class of data that can be used to assess, and possibly inform the design of new approximate XC functionals. The key benefit is a route to escape the symmetry dilemma whenever complex-polarization is relevant.

Acknowledgement This research was supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We thank Fionn Malone, Yuezhi Mao, and Narbe Mardirossian for helpful comments. The raw data for TS12 (Tables I, II, III) is available in the Supplemental Material.
It turns out that for the atoms in TS12 (i.e., C, O, S, and Si) some subtlety arises. The singlet ground state of these atoms is five-fold degenerate ($^1D$). The cRKS density is not spherically symmetric since it cannot describe this ground state degeneracy. However, we were able to obtain a non-symmetric density from exact wavefunction calculations (full configuration interaction) that is very similar to the cRKS density. If densities from these five degenerate solutions are averaged with equal weights, one obtains a spherically symmetric density. We conclude that the cRKS density is still qualitatively correct up to the ground state degeneracy.