UC Berkeley

UC Berkeley Previously Published Works

Title

Regularized orbital-optimized second-order perturbation theory.

Permalink

<https://escholarship.org/uc/item/5mp5p192>

Journal

The Journal of chemical physics, 139(24)

ISSN

0021-9606

Authors

Stück, David
Head-Gordon, Martin

Publication Date

2013-12-01

DOI

10.1063/1.4851816

Peer reviewed

Regularized Orbital Optimized Second Order Perturbation Theory

David Stück and Martin Head-Gordon

*Department of Chemistry, University of California, Berkeley,
California 94720*

(Dated: 3 December 2013)

Orbital optimized second order perturbation theory (OOMP2) optimizes the zeroth order wave function in the presence of correlations, removing the dependence of the method on Hartree–Fock orbitals. This is particularly important for systems where mean field orbitals spin contaminate to artificially lower the zeroth order energy such as open shell molecules, highly conjugated systems, and organometallic compounds. Unfortunately, the promise of OOMP2 is hampered by the possibility of solutions being drawn into divergences, which can occur during the optimization procedure if HOMO and LUMO energies approach degeneracy. In this work, we regularize these divergences through the simple addition of a level shift parameter to the denominator of the MP2 amplitudes. We find that a large level shift parameter of $400 \text{ m}E_h$ removes divergent behavior while also improving the overall accuracy of the method for atomization energies, barrier heights, intermolecular interactions, radical stabilization energies, and metal binding energies.

I. INTRODUCTION

The simplest wave function ansatz that satisfies Fermi statistics is an antisymmetrized product of single electron orbitals: the Slater determinant. The Hartree–Fock (HF) method—defined by variationally minimizing the expectation value of the Hamiltonian within the Slater determinant ansatz—gives a mean-field description that obtains approximately 99% of the total energy but ultimately fails at describing most chemical processes, such as reaction energies, with any reasonable accuracy¹. Excepting cases of strong/static correlation, where multi-reference methods are required to properly describe the physics, a primary assumption in electronic structure theory is that the HF method is a good zeroth order approximation for the true wave function. Møller–Plesset perturbation theory (MP), configuration interaction, and coupled cluster (CC) theories typically use HF orbitals as the starting point for building up to a more accurate wave function.

Unfortunately, the assumption that HF orbitals are a good zeroth order approximation does fail, particularly in the case of significant spin contamination. Restricted HF (RHF) follows our chemical intuition that electrons are paired by requiring alpha and beta electrons to have the same spatial orbitals. Removing this requirement leads to unrestricted HF (UHF)², which allows for extra variational degrees of freedom that potentially lower the energy but lead to other unintended consequences. The most clear implication of this unrestriction of the wavefunction is the introduction of spin contamination, indicating that the wavefunction is no longer an eigenfunction of the spin squared operator³. While one might not be too concerned about getting the total spin of the wavefunction correct since the energy has no direct dependence on spin degrees of freedom, such broken symmetry solutions typically lead to very poor zeroth order wavefunctions for MP2 and CC theories^{4–7}.

The point is made quite clearly in the dissociation of the lithium dimer where

MP2 using UHF orbitals dissociates correctly but gives a very poor description of the ground state potential, while RHF orbitals lead to the correct equilibrium behavior but dissociate wildly incorrectly (Figure 1). Artificial spin symmetry breaking occurs on a larger scale in C_{60} , leading UMP2 to fail at describing properties such as the single-triplet gap⁸. Spin contamination also leads to total delocalization of solitons in neutral polyenyl chains that experimentally are known to be localized over about 18 carbon atoms⁹. These examples illustrate the need for post-HF methods that are not tied to HF orbitals.

One such option is approximating the Brueckner orbitals—defined as the orbitals that give a ground state determinant with maximal overlap with the exact wave function or equivalently, the orbitals for which there are no single excitations in full configuration interaction (FCI) wavefunction¹⁰. Within the framework of FCI, one can calculate the Brueckner orbitals using either a projective or variational approach¹¹. In the projective approach, one would calculate singles amplitudes as a function of orbital rotations and minimize them to zero. For the variational approach, we constrain the singles amplitudes to be zero and minimize the CI energy with respect to orbitals, which gives the Brueckner orbitals (since they satisfy the constraint by construction). While these two approaches are identical in the FCI limit, introducing truncations to the CI expansion will lead to two distinct methods.

These Brueckner orbital methods can also be applied to the size-consistent exponential ansatz of coupled-cluster approximations, although they are not exact in the full limit¹². The projective approach has been implemented for CCSD as Brueckner doubles (BD)^{13,14} with the variational approach implemented for CCSD as orbital optimized doubles (OD)¹⁵ and for Møller Plesset theory as orbital optimized MP2 (OOMP2)¹⁶ and OOMP3¹⁷. While minimizing a non-variational method may sound troubling, the value being minimized is actually a constrained energy where the single’s contribution is set to zero, penalizing orbital solutions which have large singles

contributions to the full energy.

Many of the failings of HF orbitals can be mitigated by the use of approximate Brueckner orbital methods. Spin contamination is generally removed or significantly reduced leading to spin eigenfunctions without resorting to restricted constraints^{8,11,16,18}. In addition to rectifying spin properties, the use of approximate Brueckner orbitals has been shown to improve the description of bond lengths, frequencies, and relative energies of open shell systems^{11,15,16,18–20}. The use of the variational approach garners additional benefits due to the fact that the energy is made stable to orbital rotations. This fact gives rise to a Hellmann-Feynmann condition, simplifying response properties of the wavefunction and removing first derivative discontinuities present in UMP2 at the unrestriction point²¹.

MP2 is the one of the simplest computational methods to account for electron correlation and naturally includes long-range dispersion interactions²². It is *ab initio*, systematically improvable, and an important alternative to the more commonly used density functional theory (DFT) in cases where DFT self-interaction error is present²³. For these reasons, OOMP2 is an important method to accurately model large, open-shell systems such as radicals or organometallic compounds, striking a balance between the speed of HF and the accuracy of CCSD. For the case of Li_2 dissociation, it connects the spin pure equilibrium description to the unrestricted asymptotic limit as seen in Figure 1. Recently, variants of OOMP2 have been proposed that use a Thouless expansion representation of OOMP2²⁴ along with another approach to orbital optimization based on a one particle operator approximation to the MP2 energy²⁵.

While enabling many improvements to traditional MP2, OOMP2 brings with it the looming concern of energy divergence. Inherent in Rayleigh-Schrödinger perturbation theory is the divergence for zeroth order states with nearly degenerate energies, which translates in MP2 to divergence as the HOMO-LUMO gap goes to

FIG. 1. Li_2 dissociation curve for MP2 using restricted and unrestricted orbitals and for OOMP2 with a cc-pVDZ basis. RMP2 dissociates incorrectly and UMP2 distorts the equilibrium description while OOMP2 gets the best of both worlds by continuously connecting the two regimes, albeit with a kink due to a slight discontinuous change to the orbitals upon unrestriction.

zero. While there are classes of post-Kohn Sham theories that can properly describe small band gap systems such as RPA^{26,27} and GW theory^{28,29}, in small molecular systems such a degeneracy in the HF orbital energies is a key indicator of the presence of static correlation requiring multireference techniques rather than MP2.

For OOMP2, on the other hand, these divergences do not need to be present in the orbitals used to calculate the final energy to cause problems; due to the non-variational nature of the constrained energy, the method must only come across one of these mathematical artifacts during the optimization procedure to keep from finding a truly stable set of orbitals. In this case, by removing divergences, one could properly converge to a set of orbitals that do not contain orbital degeneracies. Thus, unlike standard MP2, the divergences present in OOMP2 can occur in many more situations since they may arise during the optimization process even if they would have not appeared in the final energy expression.

There have been many approaches taken to regularize standard MP2 theory for nearly degenerate zeroth order energies. Some apply methods of pseudo-degenerate perturbation theory where zeroth order subspaces are defined for applying perturbation theory followed by diagonalization³⁰. Another approach is to treat each second order excited state contribution as uncoupled from all others and diagonalize as in degeneracy-corrected perturbation theory (DCPT2)^{31,32} or a generalized iterative approach to diagonalize a dressed Hamiltonian³³. There have been many methods based on the repartitioning of the diagonal portion of the zeroth and first order Hamiltonian through level shifts, which leaves the energy unchanged up to first order but modifies higher order terms. One partitioning is to shift the degeneracy into the imaginary plane through a complex level shift parameter to damp out divergences^{34,35}. Other repartitioning approaches have focused on the convergence of the MP series^{36,37} and making low levels of theory stable to difficult correlations in single reference^{38,39} and multireference perturbation theory^{40,41}. In the context of complete active space second-order perturbation theory (CASPT2), a method has been developed to add a state independent level shift and then add a correction to remove the effect of the shift on the energy⁴².

These approaches all have established merits, but are more complicated than the very simplest possibility, which is the introduction of a static level shift, which perhaps surprisingly, has not been carefully explored hitherto, to our knowledge. The simple addition of a single, state independent level shift is well suited to our situation since we do not intend to properly describe these degenerate cases, but simply remove them as minima from our orbital optimization space. It is a key point that we are not trying to develop a pseudo-degenerate perturbation theory but simply modify our OOMP2 energy functional in a way that avoids artificial minima.

Accordingly, the purpose of this paper is to explore a modified OOMP2 theory with a level-shift parameter to regularize divergences that can arise during orbital

optimization. We regard this parameter as potentially serving two purposes. First, regularization itself, and, second, since stability improvements should be related to accuracy improvements, the level shift parameter is also a degree of freedom with which to remove some of the systematic error of OOMP2. After discussing the theory, we investigate the magnitude of the level shift needed for regularization, and then explore how compatible (or incompatible) it is with training the level shift parameter to remove systematic OOMP2 errors in calculated atomization energies. The transferability of the optimized parameter is then further investigated on a range of other relative energies, and also on optimized bond lengths and harmonic vibrational frequencies for molecules that are sensitive to symmetry breaking in the MP2 wavefunction.

II. THEORY

Using spin-orbital notation, the resolution of the identity⁴³ (RI) MP2 energy is given by

$$E_{\text{RIMP2}} = \frac{1}{4} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} (ia||jb)_{RI} T_{ij}^{ab}$$

where

$$\begin{aligned} T_{ij}^{ab} &= \frac{(ia||jb)_{RI}}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \\ (ia||jb)_{RI} &= \sum_{PQ}^{\text{AUX}} C_{ia}^P(P|Q)C_{jb}^Q - C_{ib}^P(P|Q)C_{ja}^Q \\ C_{ia}^P &= \sum_R^{\text{AUX}} (ia|R)(R|P)^{-1} \end{aligned}$$

This is the standard MP2 energy expression where ϵ_p are given by diagonal elements of the pseudocanonical (block diagonalized) Fock matrix and the two electron

integrals have been expanded using the resolution of the identity. For simplicity, we assume all occupied orbitals are correlated. As usual, T_{ij}^{ab} is the coefficient of the double excitation $i \rightarrow a, j \rightarrow b$ of the first order wavefunction. A subtlety to this equation is that the singles energy, which for Hartree Fock orbitals is strictly zero, is neglected even for non-HF orbitals when using OOMP2 as in OD. As mentioned above, the purpose of neglecting the singles contribution and minimizing the energy is to reach approximate Brueckner orbitals.

To minimize the energy we need the electronic gradient which is expressed as

$$\frac{\partial E}{\partial \theta_{ai}} = \sum_j \sum_b \frac{\partial U_{bj}}{\partial \theta_{ai}} (2F_{bj} + 2L_{bj})$$

with,

$$\begin{aligned} L_{ai} = & \sum_{jk} P_{jk}^{(2)} A_{aijk} + \sum_{bc} P_{bc}^{(2)} A_{aibc} \\ & - \sum_{jk} \sum_b T_{jk}^{ab} (ij||bk)_{RI} \\ & + \sum_j \sum_{bc} T_{ij}^{bc} (ab||jc)_{RI} \\ & + \sum_j F_{aj} P_{ji}^{(2)} + \sum_b P_{ab}^{(2)} F_{(bi)} \end{aligned}$$

and,

$$\begin{aligned} P_{ij}^{(2)} &= \frac{-1}{2} \sum_k \sum_{ab} T_{ik}^{ab} T_{jk}^{ab} \\ P_{ab}^{(2)} &= \frac{1}{2} \sum_{ij} \sum_c T_{ij}^{ac} T_{ij}^{bc} \\ A_{pqrs} &= (pq||rs) + (pq||sr) \\ &= 2(pq|rs) - (pr|ps) - (ps|qr) \end{aligned}$$

Standard notation is used, where F_{pq} are Fock matrix elements, $P_{pq}^{(2)}$ are elements of the correction to the two particle density matrix, and A_{pqrs} is from the HF orbital Hessian. Note that the last two terms of the Lagrangian (L_{ai}) come from off-diagonal Fock matrix elements and appear since we are not using HF orbitals.

Our proposal is to tame the divergence of the OOMP2 energy by modifying the T amplitudes which contain the energy denominators. The simplest place to start is to add a level shift to the zeroth order energies which takes the form of a small constant factor to the denominator, thus setting a lower limit to the divergence. Our new amplitudes are thus expressed,

$$T_{ij}^{ab}(\delta) = \frac{(ia||jb)_{RI}}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b - \delta}$$

This choice gives the added benefit of leaving the gradient equations unchanged except for the replacement of T with $T(\delta)$.

The level shift can be theoretically justified as a repartitioning of the zeroth order Hamiltonian as,

$$\mathbf{H}_0^\delta = \mathbf{H}_0 + \delta \cdot \mathbf{1}$$

$$\mathbf{V}^\delta = \mathbf{V} - \delta \cdot \mathbf{1}$$

which leaves the first order energy unchanged, but modifies the first order amplitudes as $T_{ij}^{ab}(\delta)$.

Another way to derive δ -OOMP2 is to start from the Hylleraas functional⁴⁴ and penalize large amplitudes by including a third term:

$$J_H(\mathbf{T}) = 2\mathbf{T}^\dagger \mathbf{V} - \mathbf{T}^\dagger (\mathbf{H}_0 - E_0) \mathbf{T} + \delta \cdot \mathbf{T}^\dagger \mathbf{T}$$

From here we minimize J_H by differentiating with respect to $\mathbf{T}$ and setting equal to zero:

$$\frac{\partial J_H}{\partial \mathbf{T}} = 2\mathbf{V} + 2(E_0 + \delta - \mathbf{H}_0)\mathbf{T} = 0$$

and we get,

$$\mathbf{T} = \frac{\mathbf{V}}{(E_0 + \delta - \mathbf{H}_0)}$$

Thus we arrive at the same equations by viewing δ as a level shift from repartitioning the Hamiltonian, or as a quadratic penalty function applied to the $\mathbf{T}$ amplitudes. More complicated (i.e. more non-linear) penalty functions are also possible⁴⁵.

III. RESULTS AND DISCUSSION

A. Divergence

While the possibility of the OOMP2 energy diverging is clear, what is unclear is under what circumstances these divergences will interfere with the optimization procedure. We are limited in our understanding of the energy as a function of orbitals due to the high dimensionality of the problem. One exception, however, is the case of H_2 in a minimal basis (in an unrestricted framework), for which the only degrees of freedom to which the energy is not invariant are the 2 rotations between occupied and virtual alpha and beta orbitals. Thus, for a given bond length, we can plot the OOMP2 energy landscape in three dimensions.

Figures 2 and 3 plot the energy surface at equilibrium and stretched geometries as a function of Given's rotations between occupied and virtual orbitals in α and β subspaces. There are a few points of note to help orient oneself in these plots. First, the $(0^\circ, 0^\circ)$ point corresponds to the orbitals obtained by diagonalizing the core Hamiltonian and, for minimal basis H_2 , corresponds quite nearly to RHF orbitals.

Second, since the two axes correspond to mixing α and β orbitals independently, all points that lie off the central diagonal ($\theta_\alpha = \theta_\beta$) will correspond to spin contaminated (unrestricted) orbitals. The final point to note is that rotation by 180° corresponds to multiplying the molecular orbital by -1 and leaves the energy unchanged. Thus the plotted region contains points corresponding to the same orbitals, but we leave this degeneracy in the plot to get a clearer visual representation of the surface.

When we look at the energy surface at equilibrium, there appears to be a clear, single minimum corresponding to the RHF solution at the origin. Although some “dents” appear near the top of the curve, we can safely say that they will have no effect on any optimization procedure since they appear near the top of a nearly 1000 kcal/mol high maximum. We can feel relatively sure that optimization on this energy surface will yield the global minimum solution without much difficulty.

FIG. 2. Dependence of the OOMP2 energy (the standard RIMP2 energy without singles contribution) on the two occupied-virtual mixing angles for the hydrogen molecule in the STO-3G basis at 0.74 Å. The region around the RHF minimum at $(0^\circ, 0^\circ)$ is well behaved.

Once we stretch the bond to 4.0 Å, we get a qualitatively different picture. Now, the new unrestricted solution shows up as a wide minimum around the point $(140^\circ, 40^\circ)$. Unlike the equilibrium case, there are points on the orbital surface where the energy diverges due to the coalescence of the HOMO and LUMO energies. In

FIG. 3. Dependence of the OOMP2 energy on the two occupied-virtual mixing angles for the hydrogen molecule in the STO-3G basis at 4.0 Å. Divergences appear for orbitals with unfavorable HF energies but very large negative MP2 energy due to HOMO-LUMO energy coalescence. There is a stable minimum near the UHF solution around $(140^\circ, 40^\circ)$, but it is not the global minimum due to the divergences.

fact, the restricted solution is surrounded by these divergences.

To get a sense of the size of the regularization parameter needed to be to remove these pits, Figure 4 plots the energy surface for δ -OOMP2 for δ values of 100 and 400 mE_h . It shows that in our toy case, one must go to values over 10 eV to tame the divergences of OOMP2. A value this high will certainly have a significant effect on absolute energies, but potentially less so on relative energies as we will see. In this case, the unrestricted solution has been restored as the global minimum and the divergences have nearly been reduced to saddle points. In general, we expect that level shifts of this magnitude should remove divergences as absolute minima, since there will be a large penalty from the first order energy for bringing the HOMO and LUMO orbital energies to degeneracy. Unfortunately, there may still be artificial, shallow local minima, but they will be easily identifiable by a HOMO-LUMO gap of zero.

FIG. 4. δ -OOMP2 orbital energy surface with level shifts, δ , of 100 mE_h (left) and 400 mE_h (right) for the hydrogen molecule in the STO-3G basis at 4.0 Å. The level shift of 400 mE_h has restored the solution near the UHF orbitals to be the global minimum and has removed the divergences.

The difficulty of looking at divergences that arise in OOMP2 is that it is a problem that depends on the specifics of the optimization algorithm and can be fixed by knowing the right answer ahead of time (since presumably the final set of orbitals should have a non-zero HOMO-LUMO gap). Rather than immediately fixing our parameter by how “regularized” it makes the optimization, we choose to look at the problem from a different perspective, by viewing δ as a semi-empirical parameter and testing how it can improve systematic errors in OOMP2. We can then assess how compatible (or incompatible) the two perspectives are.

B. Test Sets

We proceed by investigating the effect of the δ parameter on errors in calculated atomization energies compared to QCISD(T)⁴⁶ for the 148 small molecules of the G2 test set^{47,48} in the cc-pVTZ basis. The G2 test set is chosen as a fair testing grounds since thermochemistry of closed shell systems is definitely not the target of

OOMP2; in fact, for such systems, standard RIMP2 will likely be faster and more accurate. In this sense, we hope to parametrize δ -OOMP2 for general improvement, rather than fit it to a specific problem. The moderately sized cc-pVTZ basis is used in both reference and OOMP2 calculations with the matching auxiliary basis set for the resolution of the identity. In this way, we are not compensating for basis set incompleteness.

Figure 5 shows the root mean square error (RMSE) for δ -OOMP2 as well as the corresponding δ -RIMP2 and an RIMP2 and OOMP2 variant with directly scaled correlation energy (as has been previously applied to MP2^{49,50}). The other two methods are briefly considered here as a way to provide a fair comparison: inclusion of a semi-empirical parameter will necessarily improve the statistical errors and we want to make sure that the parameterization we are working with gives comparable improvements to other simple, singly-parameterized variants of MP2. These results show that a significantly large regularization parameter of around 300 or 400 mE_h optimally reduces the systematic errors of OOMP2 in atomization energies. By comparing to the other two methods it seems that the improvements are typical of parameterizations that reduce the correlation energy of MP2.

While the size of the optimal level shift seems surprisingly large, parameters on the same order of magnitude have been shown to reduce errors in CASPT2⁴² (although the study was more focused on removing the effects of the parameter rather than exploiting its reduction of errors). It is also important to compare to the previous scaled MP2 results for bond dissociations⁴⁹ and atomization energies⁵⁰ of very small molecules which actually suggest scaling the correlation by a value larger than one. These studies, however, are comparing to experimental values and not to a higher level theory in the same basis set and are thus accounting for basis set incompleteness of their double and triple zeta MP2 calculations which becomes the major factor in the results.

FIG. 5. RMS error on the G2 test set of atomization energies for δ -OOMP2, δ -RIMP2, and correlation scaled RIMP2 and OOMP2 as a function of the regularization parameter δ (bottom) or scaling parameter, s , given by $E_s = E_0 + sE^{(2)}$ (top).

Another interesting point to note is that δ -OOMP2 has a larger optimal value of δ compared to δ -RIMP2, which reinforces the idea that while MP2 typically over correlates, OOMP2 over correlates even more. In this context, the discussion of over-correlation applies specifically to relative energies of chemical significance. Thus while it is recognized that MP2 typically under estimates absolute correlation energies (except in some recently understood cases for heavy atoms^{51–53}), it tends to over emphasize the effects of correlation for relative energies as seen in the G2 results for scaled-RIMP2. Nonetheless it is important to note that this over-correlation is not a universal rule but is a significant trend seen quite clearly in the 148 molecules of the G2 test set. Finally it is also encouraging that the optimal value seems quite compatible with the values inferred as suitable for regularization in the previous section.

Performance of δ -OOMP2 on test sets that typically give standard MP2 trouble are shown in Fig 6⁵⁴. In addition to the G2 atomization energies, we have considered the S22 (weak interactions), RSE43 (radical stabilization energies), and BH76 (bar-

rier heights) test sets from Grimme’s GMTKN30 database⁵⁵ as well as a subset of the SRMBE12 that excludes second row transition metals. Calculations on the S22 test set are run in basis sets matching the original CCSD(T) values⁵⁶ without extrapolation while RSE43, BH76, and SRMB9 are run using a T-Q basis set extrapolation⁵⁷ to compare to reference values. Since errors in the various test sets can be orders of magnitude different, all RMS errors are plotted relative to that of RIMP2 using unrestricted orbitals that have been confirmed as local minima by running a stability analysis. In all cases the regularization parameter improves the performance with a degree of insensitivity to the parameter that is surprising but very promising.

While the S22 test set is not one particularly suited to orbital optimization, it is a sensitive and important case with systematic errors that, while not improved directly with orbital optimization, are reduced by scaling back the correlation energy. These improvements are seen across all subsets of interactions—hydrogen bonding, dispersion, and mixed—but most prominently improve the dispersion interactions.

Radical stabilization is where OOMP2 really shines and it is good to see that the level shift reduces error in these systems as well. It is important to recognize that the major failing of RIMP2 in these cases is due to the spin contamination in the reference, and RIMP2 can be improved using ROHF orbitals which are, however, not local minima in the full orbital space of spin polarized orbitals (and hence curves that smoothly separate bonds to correct fragments cannot be obtained).

Barrier heights are another case that requires balancing the description of two different types of systems, in this case ground and transition states. Here too, the largest errors come from cases where spin symmetry breaking is not present equally on either side of the reaction leading to cases with large errors; however, these systems can not be simply rescued by a restricting the reference as reducing the degrees of freedom leads to even worse errors.

The significant improvement seen in the description of single reference metal con-

taining compounds is very encouraging. These systems are better described by the orbital optimized reference, but also show great improvement with respect to the level shift.

FIG. 6. RMS errors of δ -OOMP2 relative to standard RIMP2 on various test sets. Without regularization OOMP2 performs worse than RIMP2 for the G2 and S22 test sets but a level shift of 400 mE_h improves δ -OOMP2 over RIMP2 and unregularized OOMP2 for all test sets.

C. Frequencies

Recent results⁵⁸ have shown that for a collection of small radicals, while standard MP2 fails dramatically for vibrational frequencies that involve symmetry breaking, OOMP2 systematically overestimates bond lengths and under estimates frequencies. This overestimation of bond lengths fits with our understanding that OOMP2 over correlates: as a single bond is pulled apart from equilibrium, electron correlations tend to grow stronger in the intermediate regime before they die off as the systems become separated. Thus the decrease in correlation energy from including the level shift parameter should decrease bond lengths, reducing the systematic errors.

FIG. 7. (a) Bond length errors vs. CCSD(T) of OOMP2, δ -OOMP2, and MP2 for five small radicals. (b) Harmonic frequencies plotted against CCSD(T) for the same five radicals. R^2 values for frequencies are 0.979, 0.998, and -0.003 for OOMP2, δ -OOMP2, and MP2 respectively. MP2 and reference CCSD(T) values taken from ref. 58.

Figure 7(a) shows the results with a $400 \text{ m}E_h$ level shift parameter confirming the improvement. The large scale failure of MP2 and systematic improvement of δ -OOMP2 over standard OOMP2 for frequencies is seen in figure 7(b). The improvement is particularly promising since it shows that the parameter independently chosen based on properties at the externally fixed geometries of the test sets is transferable to describing the correct equilibrium placement and local environment on the potential energy surface.

IV. CONCLUSION

We have presented a simple proposal for regularizing orbital optimized MP2 for near orbital degeneracy. Comparisons to standard OOMP2 and MP2 on various test sets have shown that choosing a large nonzero value for δ not only helps the method avoid diverging to artificial minima but also improves the method's accuracy. We

have selected a roughly optimal value of $400 \text{ m}E_h$ to use as the recommended value for δ -OOMP2 based initially on thermochemistry, but which shows improvements for all of the test sets studied in this work. While the cost of δ -OOMP2 is the introduction of semi-empiricism, the benefits extend beyond improved statistical errors to include the ultimate goal of stabilizing the optimization to the presence of divergences.

Since one of the greatest drawback of OOMP2 is the computational time required to iteratively calculate the MP2 energy, we plan to apply a level shift to the iterative $O(N^4)$ orbital-optimized opposite-spin scaled second-order correlation (O2)¹⁶ to make for a more tractable method. There are also other interesting possibilities for related future work. First, there is great interest in double hybrid density functionals (DHDFs)⁵⁹ at present (for instance⁶⁰⁻⁶³), including the recent possibility that orbital optimized DHDFs⁶⁴ can offer significant advantages. Very likely the inclusion of a regularization parameter as a component of an OO-DHDF would be useful both for accuracy and stability of the resulting functional. Separately, electronic attenuation has been shown to substantially increase the accuracy of MP2 theory for non-covalent interactions in finite basis sets^{65,66}. It may be that combining regularization and attenuation will further broaden the applicability of these MP2-derived methods.

V. ACKNOWLEDGEMENTS

D. S. is supported in part by the Department of Energy Office of Science Graduate Fellowship Program (DOE SCGF), made possible in part by the American Recovery and Reinvestment Act of 2009, administered by ORISE-ORAU under contract no. DE-AC05-06OR23100. This work was also supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under contract no. DE-AC02-05CH11231. We acknowledge computational resources obtained under NSF

award CHE-1048789.

REFERENCES

- ¹C. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).
- ²J. A. Pople and R. Nesbet, J. Chem. Phys. **22**, 571 (1954).
- ³P.-O. Löwdin, Rev. Mod. Phys. **35**, 496 (1963).
- ⁴N. C. Handy, P. J. Knowles, and K. Somasundram, Theor. Chim. Acta **68**, 87 (1985).
- ⁵R. H. Nobes, J. A. Pople, L. Radom, N. C. Handy, and P. J. Knowles, Chem. Phys. Lett. **138**, 481 (1987).
- ⁶M. B. Lepetit, M. Pelissier, and J.-P. Malrieu, J. Chem. Phys. **89**, 998 (1988).
- ⁷P. M. W. Gill, J. A. Pople, L. Radom, and R. H. Nobes, J. Chem. Phys. **89**, 7307 (1988).
- ⁸D. Stück, T. A. Baker, P. Zimmerman, W. Kurlancheek, and M. Head-Gordon, J. Chem. Phys. **135**, 194306 (2011).
- ⁹W. Kurlancheek, R. C. Lochan, K. Lawler, and M. Head-Gordon, J. Chem. Phys. **136**, 054113 (2012).
- ¹⁰R. Nesbet, Phys. Rev. **109**, 1632 (1958).
- ¹¹C. D. Sherrill, A. I. Krylov, E. F. C. Byrd, and M. Head-Gordon, J. Chem. Phys. **109**, 4171 (1998).
- ¹²A. Köhn and J. Olsen, J. Chem. Phys. **122**, 84116 (2005).
- ¹³R. A. Chiles and C. E. Dykstra, J. Chem. Phys. **74**, 4544 (1981).
- ¹⁴N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, and G. W. Trucks, Chem. Phys. Lett. **164**, 185 (1989).
- ¹⁵G. E. Scuseria and H. F. Schaefer, Chem. Phys. Lett. **142**, 354 (1987).
- ¹⁶R. C. Lochan and M. Head-Gordon, J. Chem. Phys. **126**, 164101 (2007).

- ¹⁷E. Soydas and U. Bozkaya, J. Chem. Theory Comput. , 1452 (2013).
- ¹⁸U. Bozkaya, J. M. Turney, Y. Yamaguchi, H. F. Schaefer, and C. D. Sherrill, J. Chem. Phys. **135**, 104103 (2011).
- ¹⁹A. I. Krylov, C. D. Sherrill, E. F. C. Byrd, and M. Head-Gordon, J. Chem. Phys. **109**, 10669 (1998).
- ²⁰F. Neese, T. Schwabe, S. Kossmann, B. Schirmer, and S. Grimme, J. Chem. Theory Comput. **5**, 3060 (2009).
- ²¹W. Kurlancheek and M. Head-Gordon, Mol. Phys. **107**, 1223 (2009).
- ²²A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (Dover Publications, 1996) p. 466.
- ²³A. D. Dutoi and M. Head-Gordon, Chem. Phys. Lett. **422**, 230 (2006).
- ²⁴J. Šimunek and J. Noga, Mol. Phys. **111**, 1119 (2013).
- ²⁵T. N. Lan and T. Yanai, J. Chem. Phys. **138**, 224108 (2013).
- ²⁶D. Langreth and J. Perdew, Phys. Rev. B **15**, 2884 (1977).
- ²⁷F. Furche, J. Chem. Phys. **129**, 114105 (2008).
- ²⁸M. Hybertsen and S. Louie, Phys. Rev. B **34**, 5390 (1986).
- ²⁹M. Shishkin and G. Kresse, Phys. Rev. B **75**, 235102 (2007).
- ³⁰P.-O. Lowdin, J. Chem. Phys. **19**, 1396 (1951).
- ³¹M.-B. Lepetit and J.-P. Malrieu, J. Chem. Phys. **87**, 5937 (1987).
- ³²X. Assfeld, J. Almlöf, and D. G. Truhlar, Chem. Phys. Lett. **241**, 438 (1995).
- ³³M.-B. Lepetit and J.-P. Malrieu, Chem. Phys. Lett. **208**, 503 (1993).
- ³⁴P. R. Surján and A. Szabados, J. Chem. Phys. **104**, 3320 (1996).
- ³⁵A. Szabados, X. Assfeld, and P. R. Surján, Theor. Chem. Acc. **105**, 408 (2001).
- ³⁶J. Olsen, P. Jørgensen, T. Helgaker, and O. Christiansen, J. Chem. Phys. **112**, 9736 (2000).
- ³⁷J. P. Finley, J. Chem. Phys. **112**, 6997 (2000).
- ³⁸A. Szabados and P. R. Surján, Chem. Phys. Lett. **308**, 303 (1999).

- ³⁹P. R. Surján and A. Szabados, J. Chem. Phys. **112**, 4438 (2000).
- ⁴⁰H. A. Witek, Y.-K. Choe, J. P. Finley, and K. Hirao, J. Comput. Chem. **23**, 957 (2002).
- ⁴¹A. A. Granovsky, J. Chem. Phys. **134**, 214113 (2011).
- ⁴²B. O. Roos and K. Andersson, Chem. Phys. Lett. **245**, 215 (1995).
- ⁴³M. Feyereisen, G. Fitzgerald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).
- ⁴⁴T. Helgaker, J. Olsen, and P. Jorgensen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, Ltd., 2000).
- ⁴⁵K. Lawler, J. Parkhill, and M. Head-Gordon, Mol. Phys. **106**, 2309 (2008).
- ⁴⁶J. A. Pople, M. Head-Gordon, and K. Raghavachari, J. Chem. Phys. **87**, 5968 (1987).
- ⁴⁷L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. **94**, 7221 (1991).
- ⁴⁸L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **106**, 1063 (1997).
- ⁴⁹M. S. Gordon and D. G. Truhlar, J. Am. Chem. Soc. **108**, 5412 (1986).
- ⁵⁰P. L. Fast, J. Corchado, M. L. Sanchez, and D. G. Truhlar, J. Phys. Chem. A **103**, 3139 (1999).
- ⁵¹S. P. McCarthy and A. J. Thakkar, J. Chem. Phys. **134**, 044102 (2011).
- ⁵²S. P. McCarthy and A. J. Thakkar, J. Chem. Phys. **136**, 054107 (2012).
- ⁵³J.-P. Malrieu and C. Angeli, Mol. Phys. **111**, 1092 (2013).
- ⁵⁴See Supplementary Material Document No. XXXXX for mean signed error, mean absolute error, and max errors for these test sets.
- ⁵⁵L. Goerigk and S. Grimme, Phys. Chem. Chem. Phys. **13**, 6670 (2011).
- ⁵⁶P. Jurecka, J. Sponer, J. Cerný, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).

- ⁵⁷A. Halkier, T. Helgaker, P. Jorgensen, and W. Klopper, Chem. Phys. Lett. **286**, 243 (1998).
- ⁵⁸U. Bozkaya and C. D. Sherrill, J. Chem. Phys. **138**, 184103 (2013).
- ⁵⁹S. Grimme, J. Chem. Phys. **124**, 034108 (2006).
- ⁶⁰S. Kozuch, D. Gruzman, and J. M. L. Martin, J. Phys. Chem. C **114**, 20801 (2010).
- ⁶¹K. Sharkas, J. Toulouse, and A. Savin, J. Chem. Phys. **134**, 064113 (2011).
- ⁶²B. Chan and L. Radom, J. Chem. Theory Comput. **7**, 2852 (2011).
- ⁶³J.-D. Chai and S.-P. Mao, Chem. Phys. Lett. **538**, 121 (2012).
- ⁶⁴R. Peverati and M. Head-Gordon, J. Chem. Phys. **139**, 024110 (2013).
- ⁶⁵M. Goldey and M. Head-Gordon, J. Phys. Chem. Lett. **3**, 3592 (2012).
- ⁶⁶M. Goldey, A. Dutoi, and M. Head-Gordon, Phys. Chem. Chem. Phys. **15**, 15869 (2013).