Convergence behavior of multireference perturbation theory : Comparative study for the BeH_2 surface of the forced degenerate and optimized zeroth order energy partitioning technique with the traditional Möller-Plesset and Epstein-Nesbet schemes.

The James Franck Institute and the Department of Chemistry

The University of Chicago

Chicago, Illinois 60637

---

84) I feel that the title is too long and unclear, but I am not sure how best to change it. How about "$\cdots$: Comparative study for the BeH_2 surface with Möller-Plesset, Epstein-Nesbet, forced degenerate, and optimized zeroth-order-energy partitioning". We are actually comparing all four partitionings with each other.

---

85) Can we remove the phrase "The recursive equations of multireference perturbation theory are solved using a configuration based algorith". I do not think it is that important how we decided to obtain the order-by-order expansion. The rest of the sentence is fine ("to study $\cdots$).

Abstract

High (through 30th) order perturbation energies are computed for BeH_2 at geometries near $Be\rightarrow H_2$ symmetric insertion transition state. The recursive equations of multireference perturbation theory are solved using a configuration based algorithm to study the difficulties in selecting the appropriate zeroth order Hamiltonian, orbitals, orbital energies, and reference functions for the computations of smooth molecular potential energy surfaces. The origin of the perturbative divergence produced by standard Möller-Plesset and Epstein-Nesbet partitionings are analyzed using a conceptually simple two-state model constructed using one state each from the reference and complementary spaces. The optimized zeroth order partitioning scheme (OPT) for double reference space computations with configurations $1a_1^2 2a_1^2 3a_1^2$ and $1a_1^2 2a_1^2 1b_2^2$ produces a truly convergent perturbation expansion through 30th order. The OPT energies are accurate in low orders as compared to the exact (197 dimensional) solution within the basis. The forced valence orbital degeneracy partitioning method (FD) also generates a truly convergent expansion for the same double reference space calculation, with slightly poorer low order energies than the OPT scheme. The BeH_2 system facilitates the consideration of larger reference spaces (constructed using three through six orbitals) where the FD method produces highly accurate energies in low orders despite the asymptotic nature of the FD perturbation expansion at very high orders. The "delayed" perturbative divergence behavior with the FD partitioning scheme (for large reference space) is shown to occur due to the "incorrect" ordering between the zeroth order energies of some reference and complementary space levels.

I. INTRODUCTION

---

107) In the introduction, I would prefer that we mention something about the work we have done on Be. Also, I would like to see more mention of OPT partitioning. Reading the tone of the paper gives me the impression that OPT has been developed by "someone else", and FD is "our method". OPT is nothing more than a further generalization of FD. We may choose to force degeneracy to obtain better energy denominators to enhance convergence. OPT just gives more degrees of freedom to obtain, in some case, improved results with better energy denominators.

86) Below its written "The perturbative approach relies on an order by order (can we use hyphons here order-by-order) expansion in which all required types of excitation process are included," This is confusion and maybe it should be removed or clarified. Since, at a given order only certain types of excitations are included, e.g., 2nd order has single and double excitations only.

———————————————————————————————————

Accurate theories for the calculation of atomic and molecular electronic structure can be broadly classified into the variational and non-variational varieties. The non-variational approaches include the single reference configuration (SR) many-body perturbation theory (MBPT) [1–4], which is a widely used and convenient procedure for a significant range of problems. The non-variational method can be classified further into (a) perturbative and (b) non-perturbative categories that invoke different philosophies and possess different advantages. The perturbative approach relies on an order by order expansion in which all required types of excitation process are included, whereas the non-variational, non-perturbative approach (e.g., the coupled cluster method [5–7]) treats particular types of excitations to all orders. Both of these categories of non-variational methods have the important virtue of maintaining size-extensivity.

———————————————————————————————————

87) Below we mention that unrestricted perturbative formalisms have spin contamination. It may be worth mentioning that these methods are also inadequate because the convergence is frequently slow when you are dealing with a single excited determinent as a reference state: There are usually other states that are more strongly coupled and close by then for a closed shell situation near the equilibrium geometry.

88) Below it is written "Schucan and Weidenmüller [20,21] appear to show that the MR perturbation series is likely to diverge if the eigenspectra of the reference space functions overlap with that of the complementary space functions" This is a good place to mention our work with the Beryllium atom. We found that computations with P–Q space overlaps can converge.

89) We write that "Mukherjee [25,26] has resolved the size-extensivity problem for the incomplete model space theories ..." . I am not sure? Other people claim that his theory is hard to implement in practice. He was in Lund and I discussed it with him. I did not get the impression that it is a straight forward thing to do.

—————————————————————————————————

Despite its great success in treating electron correlation for many types of systems, the SR-MBPT method is often inefficient (poorly convergent or even divergent) for quasidegenerate situations or for highly open shell systems where more than one reference configuration is important. These problems frequently arise in calculating excited state energies and potential energy surfaces near transition states or bond breaking regions. The use of an unrestricted Hartree-Fock reference function overcomes a portion of the convergence problem, especially near the bond breaking regions, but the perturbed wave function is spin contaminated, and the overall accuracy becomes degraded [8–10]. Several different multireference configuration MBPT (MR-MBPT) approaches (maintaining the correct spin symmetry) [11–19] have been proposed as a natural remedy to the deficiencies of the SR-MBPT method. The fundamental idea is to construct an effective Hamiltonian whose eigenvalues coincide with a subset of the eigenvalues of the exact Hamiltonian. The first size-extensive MR-MBPT formulation of Brandow [11] employs a complete reference space (often called a complete active space). The choice of a complete active space facilitates the proof of the linked cluster theorem which guarantees that the effective Hamiltonian H_{eff} has a connected diagrammatic expansion and consequently that the eigenvalues of H_{eff}, the energies, are size-extensive. While the MR-MBPT formalism resolves several shortcomings of the SR-MBPT method, it introduces another problem that has been called the "intruder state problem". Schucan and Weidenmüller [20,21] appear to show that the MR perturbation series is likely to diverge if the eigenspectra of the reference space functions overlap with that of the complementary space functions (also called virtual space functions). One solution for avoiding this problem is to use MR-MBPT with an incomplete model space, as first developed and applied by Hose and Kaldor [22]. Unlike the complete active space theories, there had been lengthy disputes [23,24] regarding the size-extensivity of the incomplete model space theories. Mukherjee [25,26] has resolved the size-extensivity problem for the incomplete model space theories by describing a "proper size- extensive normalization".

———————————————————————————————————————

90) We write "The incomplete model space MR-MBPT may suffer from poor convergence unless an appropriate reference space is selected, even though the approach is formally size-extensive." Can we remove reference to "incomplete." This is true in general, with a "complete" or "incomplete" space.

91) Also, remove the last phrase "even though the approach is formally size-extensive", since convergence and size-extensiveness is not too much related.

———————————————————————————————————————

For all practical purpose, actual perturbative computations require the truncation of the perturbative expansion, and this truncation is only meaningful if the perturbation series either converges rapidly or, at least, converges rapidly in an asymptotic sense. The incomplete model space MR-MBPT may suffer from poor convergence unless an appropriate reference space is selected, even though the approach is formally size-extensive. This proper choice of reference space becomes important in computing smooth accurate potential energy surfaces as evidenced by the general belief that the use of a single common active space is essential for computing the potential energy surface over an interesting range of geometries. It is likewise widely believed that the MR-MBPT method is incapable of providing a rapidly convergent perturbation series for the entire range of geometries. The latter belief follows from the expectation that some reference space states may escape intruder state problems for certain ranges of geometries but become plagued by these problems for other geometries. Similar questions have been raised for the intermediate Hamiltonian method, a variant of MR-MBPT [27].

———————————————————————————————————————

92) In the next sentence we credit Paldus for demonstrating "that perturbative convergence strongly depends upon the choice of the zeroth order Hamiltonian H_0." This has been known for some time. I think you may mean the zeroth order energies. Please remove credit for Paldus and just state it as a well known fact.

93) We write "Since the choice of spatial orbital energies is, in principle, arbitrary in perturbation methods, different potentials may be invoked to construct H_0. $\cdots$" Different potentials generally change the spatial orbitals, and not just their energies. The fact that orbital energies are arbitrary has not much to do with using different potentials to define H_0. Please remove this phrase, or change it.

Paldus et al. [28–30] have demonstrated that perturbative convergence strongly depends upon the choice of the zeroth order Hamiltonian H_0, i.e., on the partitioning of the exact Hamiltonian between H_0 and the perturbation V. The two general categories of partitioning are called Möller-Plesset (MP) and Epstein-Nesbet (EN) partitionings. The MP partitioning utilizes a "sum over orbitals" treatment, whereas the EN partitioning pursues a "sum over states" formulation in constructing the zeroth order Hamiltonian H_0 and the perturbation series. Since the choice of spatial orbital energies is, in principle, arbitrary in perturbation methods, different potentials may be invoked to construct H_0, and a wide range of potentials have been chosen [31,32] with varying degrees of success.

---

94) I do not know anything about the Hoffman method. Can we remove comments about it or give me a few days to read about it.

95) What does ROHF/UMP2 stand for?

96) I will read up on Murray-Davidsons method. I know that it yields similar results as CASPT2, but it can only be applied to small systems. Therefore, there must be some important differences between CASPT2 and the Murray-Davidson method. Are you sure Murray-Davidson and CASPT2 are similar?

97) We write "Considerable theoretical and computational progress has appeared for the intermediate Hamiltonian method [27], where the optimal zeroth order energies for the intermediate subspace are determined either by a judicious shifting [38–40] or by an iterative scheme [41]. However, this shifting of orbital energies introduces an additional large diagonal perturbation. Consequently, in some cases [42] this method can only push the divergent perturbative behavior to higher orders. " I do not think this is true. With the intermediate Hamiltonian you shift the energies, but you do *not* have an additional perturbation when you do these shifts. This is a nice feature. However, it is quite difficult to formulate a size-extensive intermediate hamiltonian. This problem was mentioned to me by Heully at the Ghandia conference in April. Lets mention this deficiency instead.

———————————————————————————————

Besides the use of MR-MBPT in an incomplete model space, several other methods have been also proposed to overcome the convergence difficulties induced by the presence of intruder states. For example, the Murray-Davidson [33] approach defines the orbitals by diagonalizing the Fock operator F separately in the doubly occupied, singly occupied, and virtual spaces. The H_0 is then defined by using the resulting eigenvalues as orbitals energies. Murray and Davidson also prescribe a second scheme in which the orbital energies are modified by adding a population dependent correction. The convergence behavior of the Murray-Davidson methods is similar to that for ROHF/UMP2. Roos and co-workers [34,35] use a very similar approach and, in fact, the Murray-Davidson method for ROHF is a special case of the Roos CASPT method. The main disadvantage of the methods by Pulay [36] and Roos is that a very large set of linear equations must be solved. Hoffmann [37] obtains encouraging second order energies with scaled valence orbital energies for H_0, but the higher order convergence behavior is still unknown. Considerable theoretical and computational progress has appeared for the intermediate Hamiltonian method [27], where the optimal zeroth order energies for the intermediate subspace are determined either by a judicious shifting [38–40] or by an iterative scheme [41]. However, this shifting of orbital energies introduces an additional large diagonal perturbation. Consequently, in some cases

[42] this method can only push the divergent perturbative behavior to higher orders.

---

98) We write "Their formulation is based on a Hermitianized version of Brandow's degenerate MR-MBPT theory". We may make the effective Hamiltonian Hermitian by averaging complex congugate pairs, but this has no effect on the final result. So I prefer to leave out this comment. Mention instead the orbitals that are used, $V^{(N-1)}$ type for all the valence orbitals.

99) Below we mention the OPT partitioning with FD. I would prefer that we take a short paragraph and talk about OPT a little. " .. it chooses the zeroth order energies to obtain rapid covvergence. It has worked well for H4 and Be..."

100) Below we have " *valence* space" written twice. I prefer "*refernce* space".

101) We write "The valence space in the FD partitioning is chosen based on a trade-off between generating accurate first order energies (and hence smaller perturbation corrections) from the enlarged valence space at the expense of large diagonal matrix element V due to the forced valence orbital degeneracy condition which eventually may destroy the perturbative convergence." There are other trade-offs besides the large diagonal elements of V. I prefer that we mention some of these factors–poor energy denominators and more incorrect energy ordering–with a reference to Be where this is discussed in detail.

Freed and coworkers [43–51] have introduced a rather simple approach to tackle the intruder state problem. Their formulation is based on a Hermitianized version of Brandow's degenerate MR-MBPT theory in which valence orbital degeneracy is imposed to enlarge the otherwise problematic perturbative energy denominators. This forced degeneracy condition, however, introduces an additional perturbation that enters beginning in third order, but the method significantly improves the perturbative convergence by removing serious intruder state problems. Our recent works on the HRS (rectangular H_4 system) and the Be atom demonstrate [52,53] how the forced valence orbital degeneracy approach (FD) and the

OPT (optimized zeroth order energy) partitioning offer a decently convergent series (i.e., practical convergence) in situations where the traditional MP and EN multireference perturbative series are well known to be poorly divergent for these systems due to the presence of "intruder" states. We have also demonstrated [52,53] the utility of two-state models as a diagnostic for understanding the perturbative convergence produced by various choices of reference spaces, partitioning methods, and orbitals, as well as in suggesting remedies for unsuitable choices. However, the analysis in these two studies is based on using a double reference space, which is the minimum size to qualify as a MR-MBPT computation. On the other hand, the FD partitioning scheme is generally applied with much larger valence spaces whose greater size might be thought to be more susceptible to intruder states that affect the perturbative convergence. The valence space in the FD partitioning is chosen based on a trade-off between generating accurate first order energies (and hence smaller perturbation corrections) from the enlarged valence space at the expense of large diagonal matrix element V due to the forced valence orbital degeneracy condition which eventually may destroy the perturbative convergence. The present paper addresses the pessimistic view that the large valence space ultimately must degrade the convergence properties.

---

102) Below we have "A single reference perturbative model must treat one of these two important and strongly coupled configuration state function (CSF) as lying in the complementary (Q) space, eventually creating severe convergence problems." Please remove "eventually" and change "complementary" to "orthogonal".

103) There are some errors below like "the $a_{1g}^2 b_{1u}^2$ and $a_{1g}^2 b_{2u}^2$ CSFs are equally important". It should be "$a_g^2 b_{2u}^2$ and $a_g^2 b_{3u}^2$". There are some other errors just like this following this error.

We consider the well-known difficult case of the potential energy surface for the perpendicular C_{2v} insertion of Be (^1S) into H_2 ($X^1\Sigma_g^+$). This system has been investigated by Purvis *et al.* [54–56], Simons *et al.* [57], Freed and coworkers [58], and many other groups.

The computations model the insertion path as a straight line $r = 2.54 - 0.46R$ (a.u.), where r is the H-H distance and R is the Be to center of H_2 distance. The treatment of this model reaction path is quite complicated due to the multiconfigurational nature of the electronic wavefunctions in certain regions. Purvis *et al.* [55] have shown that the perpendicular insertion of Be into H_2 requires p–orbital participation on Be and, in particular, the promotion of Be($2s^2$) to Be($2p^2$) near the critical geometry B for R=2.75 a.u. (see Figure 1) due to the quasidegenerate nature of the 2s and 2p Be orbitals. This $2s \rightarrow 2p$ promotion changes the principal $^1A_1$ configuration along the BeH_2 reaction path from $1a_1^2 2a_1^2 3a_1^2$ to $1a_1^2 2a_1^2 1b_2^2$ as illustrated in Fig. 1. Thus, at the critical or " transition" geometry B, the "excited" configuration $1a_1^2 2a_1^2 1b_2^2$ has a larger contribution than the "ground" configuration $1a_1^2 2a_1^2 3a_1^2$ to the 1^1A_1 state (using SCF orbitals from the $1a_1^2 2a_1^2 1b_2^2$ state). A single reference perturbative model must treat one of these two important and strongly coupled configuration state function (CSF) as lying in the complementary (Q) space, eventually creating severe convergence problems. When Be is moved to R=3.0 a.u. from the center of H_2 (geometry C), the 1^1A_1 state is dominated by the configuration $1a_1^2 2a_1^2 3a_1^2$, while at geometry A (indicated on Fig. 1) the configuration $1a_1^2 2a_1^2 1b_2^2$ dominates over $1a_1^2 2a_1^2 3a_1^2$. In fact, both of these RHF reference functions are unstable since lower lying UHF solutions exist. A similar multiconfigurational character also appears in the H_4 model, where at the square planner geometry (i.e., when the separation R between the two H_2 fragments is 2.0 a.u.) the $a_{1g}^2 b_{1u}^2$ and $a_{1g}^2 b_{2u}^2$ CSFs are equally important for the 1^1A_{1g} state, but as R decreases (increases) the $a_{1g}^2 b_{2u}^2$ ($a_{1g}^2 b_{1u}^2$) CSF contributes significantly. The BeH_2 case is further complicated because the $1a_1^2 2a_1^2 3a_1 4a_1$ and $1a_1^2 2a_1^2 1b_2 2b_2$ CSFs are also very important. Thus, the complete study of the BeH_2 potential energy surface is non-trivial even for describing one molecular potential energy surface.

---

104) We write "Purvis *et al.* [55] have successfully described the 1^1A_1 state BeH_2 potential energy surface using different single configuration functions for different ranges of R in coupled cluster single and doubles (CCSD) calculations, but the approach involves an arbitrary selection of reference configuration functions". Take out the phrase " but the approach involves an arbitrary selection of reference configuration functions" I think we are just repeating what was just stated. The word "arbitrary" is also confusing here. How about making the point that different reference spaces for different ranges of R can produce a bumby potential. It is also must better to be able to use a one reference space for potentials if one wants to compute gradients. Also, for dynamics I believe many points are needed, it may be difficult to get a satisfactory potential for dynamic applications when computing potentials using more than a single reference space.

105) We write "However, they fail to explain why the double reference MBPT converges at all three geometries only for the $^3B_2$ reference SCF orbital choice, an explanation that may aid in guiding further choice of reference spaces." We should change "choice of reference spaces" to "choice of MO's". Actually, do we know why this choice of orbitals works so well? Can it be that both valence orbitals are then obtained from a $V^{(N-1)}$ potential. In some cases (H4), the same reason can be used for the convergence behavior of FD, since the orbitals are defined in this manner before forcing degeneracy.

Purvis *et al.* [55] have successfully described the 1^1A_1 state BeH_2 potential energy surface using different single configuration functions for different ranges of R in coupled cluster single and doubles (CCSD) calculations, but the approach involves an arbitrary selection of reference configuration functions. Subsequently, Bartlett *et al.* [56] note that the choice of SCF orbital from a $1a_1^2 2a_1^2 3a_1^1 1b_2^1$ ($^3B_2$) reference configuration provides a converged solution for the double reference MBPT (in which the $1a_1^2 2a_1^2 3a_1^2$ and $1a_1^2 2a_1^2 1b_2^2$ CSFs are chosen as the reference space) at the A, B and C geometries. However, they fail to explain why the double reference MBPT converges at all three geometries only for the $^3B_2$ reference SCF orbital choice, an explanation that may aid in guiding further choice of reference spaces.

---

106) We mention below that "forced degeneracy (FD) partitioning scheme which uses a more sensible selection of valence orbitals enhance the perturbative convergence in low orders and to provide a usefully asymptotic perturbation series." We should do the same for OPT partitioning: "It chooses the zeroth order energies an an optimal way to produce a rapid convergence."

———————————————————————————————————

The present work analyzes the convergence problems encountered with MP and EN partitioning for double reference MBPT computations based on the $1a_1^2 2a_1^2 3a_1^2$ $(^1A_1)$ state SCF orbitals and orbital energies. We demonstrate that a converged solutions may still be obtained for all the three geometries using the same reference configuration $1a_1^2 2a_1^2 3a_1^2$ $(^1A_1)$ SCF orbitals by (1) optimizing the zeroth order state energies, i.e, by employing our OPT method, or (2) by applying the forced degeneracy (FD) partitioning scheme which uses a more sensible selection of valence orbitals enhance the perturbative convergence in low orders and to provide a usefully asymptotic perturbation series. Interestingly, recent computations for *cis*−butadiene (third order MR-MBPT) and for CH_2 by Freed and coworkers [59,60] also demonstrate the same trends.

———————————————————————————————————

107) Please remove the last sentence of the next paragraph. "We also discuss the possibility of generating a smooth potential energy curve using different valence spaces at different geometries. "

———————————————————————————————————

Sections II and III briefly outline the multireference configuration perturbation theory and the choice of the zeroth order Hamiltonian. Section IV describes the convergence criteria for the perturbation expansion of a simple two-state model that is used in later sections to explain the behavior of the perturbation expansion with the full 197 state computation. Section V presents information concerning the basis set and the model spaces used. Section VI describes why certain partitioning schemes yield convergence problems. The source of these convergence problems is analyzed using two-state models constructed from the full problem by selecting one important state each from the reference and secondary space states (called an interspace pair of states). Section VII summarizes the results of the OPT and FD partitioning (with different valence space orbitals) along with other methods that have been used to compute the $1^1A_1 \rightarrow 2^1A_1$ vertical excitation energies of BeH_2. We also discuss the possibility of generating a smooth potential energy curve using different valence spaces at different geometries.

II. BASIC FORMALISM OF MULTIREFERENCE MANY-BODY PERTURBATION THEORY (MR-MBPT)

The exact Hamiltonian is first partitioned into

$$H = H_0 + V, \tag{2.1}$$

where H_0 is the unperturbed Hamiltonian and V is the perturbation. We assume that the Schrödinger equation for the unperturbed Hamiltonian,

$$H_0|\Phi_i\rangle = E_i^0|\Phi_i\rangle, \tag{2.2}$$

provides a complete set of eigenfunctions $|\Phi_i\rangle$ with corresponding eigenvalues E_i^0. The eigenfunctions of H_0 are then divided into two subspaces defined by the two complementary projectors P and Q, where

$$P = \sum_i^d |\Phi_i\rangle\langle\Phi_i| = \sum_i P_i, \tag{2.3}$$

and

$$Q = 1 - P. \tag{2.4}$$

The P subspace of dimension d is variously called the model, the reference, or the valence space, while its orthogonal complement Q is formed by the remaining eigenvectors. The model space functions $|\Psi_i^0\rangle$ are defined as the projections of the exact eigenfunctions $|\Psi_i\rangle$ onto the reference or model space,

$$|\Psi_i^0\rangle = P|\Psi_i\rangle. \tag{2.5}$$

Alternatively, the exact eigenfunctions $|\Psi_i\rangle$ can be reproduced from the model space function $|\Psi_i^0\rangle$ with the aid of the wave operator Ω,

$$|\Psi_k\rangle = \Omega|\Psi_k^0\rangle \tag{2.6}$$

Using these definition of P, Q and Ω, the exact and the perturbed Schrödinger equations may be cast into an effective eigenvalue equations of the form

$$H_{eff}|\Psi_i^0\rangle = E_i|\Psi_i^0\rangle, \tag{2.7}$$

where H_{eff} is given by

$$H_{eff} = PH\Omega P. \tag{2.8}$$

An order by order perturbation expansion of Ω is obtained by applying Q on the left side of to the Eq. (2.7) and by separating the Hamiltonian as in Eq. (2.1) [14,55],

$$[\Omega, H_0]P = [V\Omega - \Omega V\Omega]P. \tag{2.9}$$

Now define the perturbative expansion of the wave operator Ω as

$$\Omega = 1 + \Omega^{(1)} + \Omega^{(2)} \cdots \tag{2.10}$$

Substituting the expansion for Ω into Eq. (2.9) generates the nth order expression for Ω as

$$[\Omega^{(n)}, H_0]P = [V\Omega^{(n-1)} - \sum_{m=1}^{n-1} \Omega^{(m)} V\Omega^{(n-m-1)}]P, \tag{2.11}$$

and the corresponding matrix elements of the n-th order contribution to H_{eff} are given by

$$\langle\phi_i^0|H_{eff}^{(n)}|\phi_j^0\rangle = \langle\phi_i^0|PV\Omega^{(n-1)}P|\phi_j^0\rangle. \tag{2.12}$$

The expressions for the effective Hamiltonian matrix elements through third order is then obtained as

$$\langle\alpha|H_{eff}^{(1\to3)}|\beta\rangle = E_\alpha^0\delta_{\alpha\beta} + \langle\alpha|V|\beta\rangle + \sum_{m\in Q}\frac{\langle\alpha|V|m\rangle\langle m|V|\beta\rangle}{E_\beta^{(0)}-E_m^{(0)}} + \sum_{m,n\in Q}\frac{\langle\alpha|V|m\rangle\langle m|V|n\rangle\langle n|V|\beta\rangle}{(E_\beta^{(0)}-E_m^{(0)})(E_\beta^{(0)}-E_n^{(0)})}$$

$$- \sum_{m\in Q,\gamma\in P}\frac{\langle\alpha|V|m\rangle\langle m|V|\gamma\rangle\langle\gamma|V|\beta\rangle}{(E_\beta^{(0)}-E_m^{(0)})(E_\gamma^{(0)}-E_m^{(0)})} \tag{2.13}$$

The above effective Hamiltonian H_{eff} in Eq. (2.13) is non-Hermitian as is evident from its energy denominators. Introducing a Hermitized form of Ω enables generating a Hermitian H_{eff} [11,13], which through third order is equivalent to using $\frac{1}{2}(H_{eff}+H_{eff}^\dagger)$. Provided the dimension of the Q space is not very large, the order by order computation of H_{eff} from Eq. (2.11) is relatively straightforward and is performed in this present work.

Although not explicitly indicated in the above equations, the wave operator Ω introduced here actually depends upon the model function on which it operators, i.e., is a state or ket dependent wave operator. Further, the above derivation of H_{eff} assumes that Ω satisfies intermediate normalization (i.e., PΩP=P), which, in fact, is not mandatory [25].

III. CHOICES OF ZEROTH ORDER HAMILTONIAN H_0

The choice of the zeroth order Hamiltonian is, *in principle*, arbitrary and is, therefore, at our disposal. Several types of zeroth order Hamiltonians H_0 are considered here.

A. Möller-Plesset (MP) partitioning

————————————————————————————————————

1) Should we continue to define MP partitioning by Eq. (3.1)? Is it better to use the convention that orbitals and energies are defined by a fock operator for MP.

————————————————————————————————————

The MP partitioning uses a "sum over orbitals" form of H_0. The most general diagonal form of such a H_0 is given by

$$H_0 = \sum_{c}^{N_c} \epsilon_i a_c^\dagger a_c + \sum_{v}^{N_v} \epsilon_v a_v^\dagger a_v + \sum_{e}^{N_c} \epsilon_e a_e^\dagger a_e, \tag{3.1}$$

where N_c, N_v, and N_e are the numbers of core, valence, and excited orbitals, ϵ_c, ϵ_v, and ϵ_e are the corresponding orbital energies, and $a_i^\dagger$ and a are the usual creation and annihilation operators, respectively. The orbitals are usually defined as eigenfunctions of Fock operators. The most natural choice for the orbital energies ϵ_i are the eigenvalues of the Fock operator defining the orbitals. However, any choice of orbitals and orbital energies may, in principle, be chosen, and the objective is to find choices that assure the most rapid perturbative convergence.

B. Forced valence orbital degenerate (FD) partitioning

———————————————————————————————————

2) We need a citation here that defines the multiple Fock operators. The N2 paper is quite exhaustive in describing it. **3) I think** we should describe the Fock operators used for BeH2, as is done for Be. It is best to indicate that the high lying valence orbitals can be obtained by a unitary transformation, (also discussed in N2 paper). (Note that for Be the high lying orbitals are not IVO's just their orbital energies are defined in that way.

———————————————————————————————————

This is a variant of the MP partitioning method. Since the orbitals and the energies ϵ_i in Eq. (3.1) are, in principle, arbitrary, Freed and co-workers uses multiple Fock operators to obtain the spatial orbitals. The valence orbital energies are obtained using $V^{(N-1)}$ potentials and are then forced to be degenerate in zeroth order. The form of the diagonal H_0 in the FD partitioning is given by

$$H_0 = \sum_{c}^{N_c} \epsilon_c a_c^\dagger a_c + \bar{\epsilon}_v \sum_{v}^{N_v} a_v^\dagger a_v + \sum_{e}^{N_e} \epsilon_e a_e^\dagger a_e, \tag{3.2}$$

where the average valence orbital energy $\bar{\epsilon}_v$ is obtained from the original set of valence orbital energies by the democratic averaging

$$\bar{\epsilon}_v = \frac{\sum_i^{N_v} \epsilon_v}{N_v}. \tag{3.3}$$

The above degeneracy condition introduces a diagonal perturbation $\delta V = \epsilon_v - \bar{\epsilon}_v$ that contributes beginning in third order. The magnitude of δV directly depends upon of the spread of the original valence orbital energies $\{\epsilon_v\}$ before averaging. In fact, some third order computations with small, quasidegenerate valence spaces do not require valence orbital energy averaging.

C. Epstein-Nesbet (EN) partitioning

————————————————————————————————-

4)It is better to clearly state from the beginning which type of EN partitioning we use– Barycenter. How about using what was done for Be, since it is discussed in that paper in detail. **5) EN does** not need to be defined, as stated below, by a determinental basis. I think any basis can be used for EN (e.g. CSF basis), as long as the diagonal elements of V are zero. Of course, Barycenter–EN deviates from this definition. Some how this must be explained.

————————————————————————————————

The EN partitioning uses a " sum over states" form of H_0. The most general form of a diagonal H_0 in this partitioning scheme can be written as

$$H_0 = \sum_i |i\rangle E_i^0 \langle i|, \tag{3.4}$$

where the sum over i runs over all states and E_i^0 is the ith state zeroth order energy. The EN partitioning chooses $E_i^0 = \langle i|H|i\rangle$ where $|i\rangle$ is a determinantal state. This choice makes the diagonal elements of V vanish for a determinantal basis, but not for a configuration state functions (CSF) basis as used in the computations below. Hence, the EN partitioning is not invariant under a unitary transformation of orbitals. [The configuration state function based (CSF) EN partitioning applied here, is actually a hybrid scheme called "barycentric EN partitioning" [61], where H_0 are diagonal in determinantal basis.]

D. Optimized zeroth order (OPT) partitioning:

———

6)Lets not call it a variant of EN partitioning, since we can also optimize orbital energies. Also, the optimized zeroth order energies have nothing to do with EN. **7)I think** the sentence that begins with "The lowest zeroth order.." is difficult to understand. Can we use some of the things we wrote for the Be paper. I think it more clear there. **8)Lets** clearly indicate that the unoptimized energies are equivalent to barycenter EN, not just EN. **9)Eq. (3.6)** says that $\hat{E}_i^0$ is an optimized zeroth order energies, but only some of the $\hat{E}_i^0$ are optimized. **10)Should** we have a "hat" $\hat{.}$ on $\hat{E}_i^0$? It is not an operator.

———

This is a variant of the EN partitioning method, in which a small subset of the zeroth order state energies are determined in an optimal manner. The "optimal performance" is defined as follow: The lowest twenty zeroth order state energies are selected so that they minimize the sum of the absolute deviations of the perturbative energies through third and fourth orders from the infinite order (converged) FCI (full configuration interaction) ground state energy E_{gs} as calculated *within* the subspace of the same twenty lowest zeroth order states. More precisely, the OPT partitioning involves minimizing the sum of the third and fourth order absolute deviations,

$$|E_{gs} - E_3| + |E_{gs} - E_4|, \tag{3.5}$$

where E_3, and E_4 are the third and fourth order perturbation computations of E_{gs} as obtained using the same twenty zeroth order states. Since Eq. (3.4) is invariant to a constant shift of all twenty zeroth order state energies, for convenience, we select the ground state zeroth order energy to be fixed at its EN value. The state energies for all the remaining Q-space CSFs are also fixed at their EN values. Thus, the form of H_0 in this partitioning schemes is written as

$$H_0 = \sum_i |i\rangle \hat{E}_i^0 \langle i|, \tag{3.6}$$

where $\hat{E}_i^0$ is the optimal zeroth order energy of state $|i\rangle$.

IV. TWO STATE SYSTEMS AS AN IMPORTANT DIAGNOSTIC

———————————————————————————————————————-

11)I think we should take the clause "(defined as one state each from the P and Q space)" out of parenthesis, since it is an important definition. **12)I do** not think Eq. (4.4) is correct, since $\Delta\epsilon$ can be negative and R_c must be positive.

———————————————————————————————————————-

The convergence properties for the BeH_2 computations are explained transparently in subsequent sections by the use of simple two-dimensional models that are constructed from a pair of interspace states (defined as one state each from the P and Q space) taken from the full problem with 197 CSFs. The convergence behavior is investigated by considering the parameterized Hamiltonian $H(z)$,

$$H(z) = H_0 + zV, \tag{4.1}$$

where z is the complex perturbation parameter. The $z = 1$ limit recovers the exact two-state H, while $z = 0$ produces the unperturbed systems. Denote $|p\rangle$ and $|q\rangle$ as the P and Q space states of the two-dimensional model, while ϵ_p and ϵ_q are their zeroth order energies, respectively. The exact eigenvalues of the two-dimensional Hamiltonian $H(z)$ can be expressed as

$$E_{\pm}(z) = \frac{1}{2}\mathrm{Tr}\,\mathbf{H} \pm \frac{1}{2}\{[\Delta\epsilon - (\Delta\epsilon - \Delta H)z]^2 + 4V_{pq}^2 z^2\}^{1/2}, \tag{4.2}$$

where

$$\begin{aligned} V_{pq} &= \langle p|H(z=1)|q\rangle, \\ \Delta\epsilon &= \epsilon_q - \epsilon_p, \\ \Delta H &= \langle q|H(z=1)|q\rangle - \langle p|H(z=1)|p\rangle, \end{aligned}$$

and the z dependent trace of the Hamiltonian matrix $\mathbf{H}$ is given by

$$\mathrm{Tr}\,\mathbf{H} = \langle p|H|p\rangle + \langle q|H|q\rangle.$$

Both eigenvalues become degenerate in the complex $z-$plane at the pair of branch points z_d and z_d^*, where use of Eq. (4.2) gives

$$z_d = \frac{\Delta\epsilon}{4V_{pq}^2 + (\Delta\epsilon - \Delta H)^2}[(\Delta\epsilon - \Delta H) + 2V_{pq}i]. \tag{4.3}$$

The radius of convergence R_c for the single reference two-state Rayleigh–Schrödinger perturbation expansion is $R_c = |z_d| = |z_d^*|$ and follows from Eq. (4.3) as

$$R_c = \frac{\Delta\epsilon}{\sqrt{(\Delta\epsilon - \Delta H)^2 + 4V_{pq}^2}}. \tag{4.4}$$

To achieve a convergent perturbation series, it is necessary to have $R_c \geq 1$, which occurs only if the numerator in Eq. (4.4) exceeds the denominator. This condition for convergence implies that $\Delta\epsilon$ (the zeroth-order energy difference) must satisfy the requirements

$$\Delta\epsilon \begin{cases} \geq \frac{1}{2}[\Delta H + \frac{4V_{pq}^2}{\Delta H}] & \text{if } \Delta H > 0, \\ \leq \frac{1}{2}[\Delta H + \frac{4V_{pq}^2}{\Delta H}] & \text{if } \Delta H < 0. \end{cases} \tag{4.5}$$

Thus, both $\Delta\epsilon$ and ΔH must have the same sign; otherwise the two-state perturbation series diverges. Also, when $\Delta H \neq 0$, Eq. (4.5) indicates that $\Delta\epsilon$ can always be selected so that $R_c > 1$. Wilson *et al.* arrive at the same conclusion by examining the maximum radius of convergence R_c as a function of $\Delta\epsilon$ [62]. Note that the features exhibited by Eqs. (4.3)–(4.5) are also displayed by the general behavior for the BeH_2 computations that follow of 197 CSF. This behavior is modeled is Sec. VI by selecting the most influential pairs of interspace states and by analyzing the character of the full computation for $z \approx z_d$. Thus, two-state interspace models provide both convergence criteria and a useful diagnostic tool for assessing the probable perturbative behavior of various MR partitioning methods.

Intruder states are defined as the orthogonal (Q) space states responsible for destroying the convergence of a perturbation expansion. The presence of intruder states is usually detected by observing the variation of the eigenvalues as a function the of perturbation parameter z for real z. An avoided crossing occurs at a point z_{ac} where the two eigenvalues are closest. The avoided crossing point is computed in the two-state model by minimizing $[E_-(z) - E_+(z)]$ with respect to z,

$$z_{\rm ac} = \frac{\Delta\epsilon(\Delta\epsilon - \Delta H)}{4V_{pq}^2 + (\Delta\epsilon - \Delta H)^2}. \tag{4.6}$$

Comparing Eqs. (4.3) and (4.6) shows that $z_{\rm ac}$ is simply $\mathrm{Re}\, z_d$. (Note that the eigenvalues $E_-(z)$ and $E_+(z)$ cannot be degenerate for real z, unless $V_{pq} = 0$). If the avoided crossing appears for $0 < z_{\rm ac} < 1$ and the perturbation expansion diverges ($R_c < 1$), the convergence is disrupted by what is termed a "front door" intruder state. Similarly, an avoided crossing for $z_{\rm ac} < 0$ is called a back door intruder state [63,64] when $R_c < 1$. Equation (4.5) implies that a back door intruder state $z_{\rm ac} < 0$ occurs whenever ΔH and $\Delta\epsilon$ have the same signs and satisfy $|\Delta H| > |\Delta\epsilon|$. On the other hand, if $\Delta\epsilon$ and ΔH are of opposite signs (called incorrect energy ordering) or if $|\Delta H| < |\Delta\epsilon|$, then the avoided crossing appears for $z_{\rm ac} > 0$.

---

13) $\frac{1}{2}\mathrm{Tr}\,\mathbf{H} = 0$ does not occur "when the z-dependent diagonal elements of H are degenerate"! So below I took it out in the sentence following Eq. (4.7).

When incorrect energy ordering is present in the two-state system, the expansion is always divergent ($R_c < 1$). Also, with incorrect ordering present, the limit of $V \to 0$ has $\mathrm{Im}\, z \to 0$, and z_d is real and positive with

$$z_d \to \frac{\Delta\epsilon}{(\Delta\epsilon - \Delta H)}. \tag{4.7}$$

Equation (4.7) also emerges when the z-dependent diagonal elements of H are degenerate. Thus, in actual calculations involving a large number of states, if an incorrectly ordered pair of interspace states are weakly coupled, both among themselves and to other states, then both z_d and $z_{\rm ac}$ are approximately equal to z, where the diagonal elements of $H(z)$ are degenerate for the interspace states. Furthermore, such cases yield a very pronounced avoided crossing since the degeneracy point z_d occurs close to the real axis. This situation is found in Sect. V for the FD partitioning computations with a three-orbital reference space.

By contrast, when a two-state system has the zeroth order states correctly ordered, the limit of $V \to 0$ then produces $R_c \to \infty$ corresponding to a convergent expansion.

———————————————————————————————————————-

14)We must make some modification for the following statement on EN partitioning. We should indicated that this is not true for Barycenter EN but is so for "normal" EN using any basis, determinantal of CSF.

———————————————————————————————————————-

In addition, the special case of $\Delta H = \Delta\epsilon$ (Epstein–Nesbet partitioning for a determinantal basis) yields the avoided crossing at $z_{\rm ac} = 0$ since equation (4.3) implies that this case yields z_d as purely imaginary.

———————————————————————————————————————-

15)Below we write "in general, we neither know nor care whether the perturbation series converges". We should care if it is convergent. Why are we then analyzing the high order convergence. We may not always recover all the correlational energy needed for an assymptotic series. Also, our paper is of great interest to people doing MRCC. The divergent behavior of MRCC can be identical to the order-by-order results we have, except that we include all excitations, and MRCC neglects the higher-body ones, this is especially true when MRCC is solved in an iterative fashion corresponing to an order-by-order perturbative expansion. So please make a minor modification of the next sentance. You can retain the basic point.

———————————————————————————————————————-

It should be emphasized that, in general, we neither know nor care whether the perturbation series converges, since this information has a little practical value for computations truncated at low orders. Mathematical convergence depends upon the behavior of the indefinitely high order terms, whereas low order truncations are often very accurately provided by formally divergent asymptotic series. For example, the power series expansion of the Bessel function

$$J_0(\epsilon) = 1 - \frac{1}{4}\epsilon^2 + \frac{1}{64}\epsilon^4 - \frac{1}{2304}\epsilon^6 + ... \quad (4.8)$$

has an infinite radius of convergence. However, unless ϵ is small, many terms are required to achieve accurate results from Eq. (4.8). On the other hand, the asymptotic expansion (for $x \rightarrow 0$)

$$J_0(\frac{1}{x}) \approx \sqrt{\frac{2x}{\pi}}[(1 - \frac{9}{128}x^2 + ..)cos(\frac{1}{x} - \frac{\pi}{4}) + (\frac{1}{8}x - \frac{75}{1024}x^3 + ...)sin(\frac{1}{x} - \frac{\pi}{4})] \quad (4.9)$$

is divergent no matter how small x is, but the first few terms provide good accuracy for moderately small x. Likewise, many formally divergent series are asymptotically convergent and are quite useful when truncated at low orders (e.g., Stirling's formulas for the expansion of $\ln n!$ and $\sin(2\epsilon)$). Thus, when divergent series are obtained from the full scale calculations, these series are analyzed below to assess their useful asymptotic character and their behavior when resumed by Padé approximants.

V. BASIS SET AND MODEL SPACE

———————————————————————————————-

16)We should mention from the onset that the $1a_1$ orbital is frozen. **17)Are** you sure that Bartlett et. al. has develped the algorithm. You really mean the theory, and not the program. Even the theory, are we sure Bartlett did it first? **18)the** geometries A B and C are not properly defined. What is the H–H bond distance? Should we give a citation here where A, B and C are define?

———————————————————————————————

We employ a contracted [12s3p/3s1p] set of Gaussian functions for Be and [4s/2s] contracted Gaussian functions (with scaling factor of 1.2) for the H atoms. This basis is the same as applied by Purvis and Bartlett in their ground state computation for BeH_2 [55]. Ground state SCF calculations produces the minimum energy state as a $1a_1^2 2a_1^2 3a_1^1 1b_2^1\ {}^3B_2$ state for all the geometries considered here. The next two low lying states,using the same set of orbitals, have the configurations (symmetries) of $1a_1^2 2a_1^2 3a_1^2$ $(1\,{}^1A_1)$ and $1a_1^2 2a_1^2 1b_2^2$ $(2\,{}^1A_1)$

and cross near geometry B (see Fig. 1). Our first reference space for all the geometries contains the two zeroth order $1a_1^2 2a_1^2 3a_1^2$ and $1a_1^2 2a_1^2 1b_2^2$ states. This double reference space is used to study the convergence behavior for the two lowest $^1A_1$ excited states with all four different partitioning schemes, the MP, EN, FD and OPT partitionings. The double reference space calculation for the states of $^1A_1$ symmetry yields a $Q-$space with 195 states since the $1a_1$ orbital is kept frozen. We compute through 30th order the order-by-order state energies and vertical excitation energy for the $1^1A_1 \rightarrow 2^1A_1$ transition using the configuration based MR-MBPT algorithm developed by Bartlett et al. [55]. Excitation energies for all reference spaces are taken relative to the 1^1A_1 state, although the $^3B_2$ state is the lowest lying SCF state. We also compute the order by order state energies using FD partitioning with larger reference spaces in order to analyze how the convergence behavior is affected by enlarging the reference space.

——————————————————————————————-

19)Below we discuss the trade-offs present when increasing the reference space. We should mention that there are various factors that may adversely affect the convergence when the reference space in enlarged, not just the additional perturbation. These factors are discussed in the Be paper. We should cite this paper and mention the other factors. **20)Should** we also define in this section the larger reference spaces? This must be added somewhere.

——————————————————————————————-

Since increasing size of the reference space enhances the diagonal perturbation matrix elements (due to the reduced quasidegeneracy among the valence orbitals), it might be expected that the perturbation series diverges more quickly with the larger than the smaller reference spaces. On the other hand, the larger reference space provides improved first order energies and smaller overall perturbation corrections. The question we address with the larger reference spaces is which of these competing factors dominates when larger reference space computations are truncated at low (e.g., third) orders.

——————————————————————————————-

21)We say that MP and FD are different, but we have defined FD as a variant of MP. This will produce much confusion. **22)What version** of MP are we doing? How are the orbitals defined for MP? Is it an open shell restricted ground state? **23)We should** mention that the IVO computations are actually not an SCF, but are obtained from a unitary transformation using a f^{N-1} Fock operator (see N2 paper). This Fock operator has occupation $1a_1^2 2a_1^2 3a_1^1$. This is much simpler, but equivalent. **24)Are we** defining the FD orbitals in the same way for all geometries: We have made some arbitrary choice regarding the first SCF calcualation. Usually we choose the ground state. Here we are selecting an excited state $1a_1^2 2a_1^2 3a_1^2$, and this state is not always the lowest excited state. **25)Why we** do it in this manner should be discussed. Also this section has the title "Basis set and Model space". **26)Can we** use reference space instead? **27)Should we** put the definition of the orbitals in section III when FD partitioning is defined. The title of this section does not really indicate that we are going to define the MOs for FD partitioning.

Both the orbitals and orbital energies are chosen somewhat differently with the FD partitioning scheme than with conventional MP methods. We illustrates these differences using the double reference space FD-MBPT computation with the $1a_1^2 2a_1^2 3a_1^2$ and $1a_1^2 2a_1^2 1b_2^2$ CSFs as the reference space. The $1a_1$, $2a_1$ and $3a_1$ orbitals and orbital energies are first obtained from a closed shell SCF computation for the $1a_1^2 2a_1^2 3a_1^2$ state. The $1b_2$ unoccupied valence orbital and orbital energy are then computed as improved virtual orbital (IVO) from a $1a_1^2 2a_1^2 3a_1^1 1b_2^1$ state SCF in which only the $1b_2$ is permitted to vary. This procedure yields a $1b_2$ orbital that is more suited to describe excited states than the virtual $1b_2$ orbital from the $1a_1^2 2a_1^2 3a_1^2$ state SCF, and the $1b_2$ orbital energy is lower than that of the virtual $1b_2$ orbital, making the FD reference space more quasidegenerate than the conventional MP choice. The excited FD partitioning orbitals and orbital energies are then obtained by diagonalizing the $1a_1^2 2a_1^2 3a_1^2$ state Fock operator with the core and valence orbitals frozen. The $3a_1$ and $1b_2$ orbital energies are then replaced by their arithmetic average to impose degeneracy on the valence space. A similar procedure is followed for the larger reference space. Those valence orbitals not occupied in $1a_1^2 2a_1^2 3a_1^2$ state are defined as IVOs, and eigenvalues of the IVO

Fock operators provide the corresponding orbital energies before averaging.

VI. CONVERGENCE PROBLEM FOR BERYLLIUM HYDRIDE WITH DIFFERENT PARTITIONING SCHEMES

A. Convergence difficulties with Möller-Plesset (MP) and Epstein-Nesbet (EN) partitioning schemes

Given the choice of the $|1a_1^2 2a_1^2 3a_1^2\rangle$ and $|1a_1^2 2a_1^2 1b_2^2\rangle$ CSFs as the model or reference space states, the computed MR-MBPT energies for the lowest two $^1A_1$ states are displayed in Tables I, II and III for geometries A, B and C, respectively, and for the EN, MP, FD and OPT partitioning schemes. The tables demonstrate the previously known fact [56] that the MP and EN partitioning multireference perturbation expansions suffer from convergence problems for the both $^1A_1$ states of BeH_2 and for all the geometries, except geometry A with MP partitioning. The analysis below demonstrates that the severe divergence encountered with the MP and EN partitionings arise from the presence of intruder states which render these series divergent. Detailed analysis is only presented for the "transition state " geometry B.

———————————————————————————————

28)Are we sure about the next sentence?

———————————————————————————————

The convergence difficulties at the other two geometries can be shown to occur for similar reasons.

———————————————————————————————

29)Can we use an *italics* to represent are energies. $[\mathrm{E_m}(z) - \mathrm{E}\bar{(}z)]$ looks kind of ugly. How about $[E_m(z) - \bar{E}(z)]$. Also note that the bar on $\bar{E}(z)$ must be over the E and not like $E\bar{(}z)$. **30)So** please fix it. These changes also need to be put into the figures. If you have problems let me know. Gnuplot can be used with LaTeX to get italic symbols. **31)The**

next two sentences start with "Figure 2". Please remove this repitition. **32)State** that we are examining Geometry B. **33)Can we** try a different scaling. The range is quite large. **34)"m-th" should** be mth, m is in italics. **35)Figure 2 is** very cluttered. **36)Can we** remove the fouth state. It is not needed. **37)Use a** smaller font for the zeroth-order labels on the figure, and indicate that the z-dependent states are labeled by the state that they correspond to at z=0. This has not been mentioned. **38)It** is necessary to define what the radius of convergence is: It is related to where the second and third excited states become degenerate in the complex plane. Then, the reader knows what we are looking for. You mention below that there is an interspace degeneracy. However, we have defined interspace states only as the zeroth order states, when z=0 only. **39a)We need** to introduce the idea of one-to-one correspondance between the zeroth order states and the z-dependent states. **39b)Thus,** lets say that the avoided crossing is from two states that both correspond to P-space states when z goes to zero. **40)We also** say that the avoided crossing presents no impediment. This is confusing jargon that should be removed. An avoided crossing can never impede convergence. It can only indicate approximately what the radius of convergence is. For H4 we also found an avoided crossing for FD involving states that correspond to the P space when z goes to zero. **41)Is the** same thing happening here as H4. Is there a degeneracy in the expecatation values of H for 2 zeroth order states. **42)Can we** indicate P–corresponding states with one kind of line, say solid, and Q–corresponding with, say a dotted line.

————————————————————————————

Figure 2 depicts the exact $^1A_1$ eigenspectrum for the four lowest states with MP partitioning as function real z (with Im $z = 0$). Figure 2 presents $[\mathrm{E_m}(z) - \bar{\mathrm{E}}(z)]$ for Re z, where $\mathrm{E_m}(z)$ is the z-dependent eigenvalue of the mth $^1A_1$ state and $\bar{\mathrm{E}}(z)$ is the average of all the four state energies. The advantage of depicting $[\mathrm{E_m}(z) - \bar{\mathrm{E}}(z)]$ over $\mathrm{E_m}(z)$ is that the former can provide a smaller range for the abscissa than the later. Note that the FCI energies ($z = 1$) are independent of the partitioning schemes, but the z-dependent FCI eigenvalues of $H(z)$ are functions of z. The two evident avoided crossings for $z \approx 0.8$ between the 1^1A_1

and 2^1A_1 and between the 3^1A_1 and 4^1A_1 states z-dependent energies in Fig. 2 present no impediment to the double reference perturbative convergence since the crossings involve intraspace (within P and within Q spaces, respectively) avoided crossings.

———————————————————————————————

43)How about end the paragraph here. **44)Maybe we** should remove the "Figure 3" repition in the first 2 sentences. **45)Below we** use the word "correlation". I prefer the word "corresponding" to describe what happens to states when z goes to zero. This convention was used for Be. Lets stick to it so people can easily read both paper. I do not like that we are presenting energy differences in the figures. **46)Can we** just give the energies to show where the states are degenerate. This is the way it was done before. We may need two graphs, one for the complex and one for the real energies, but this is no problem. One should be able to see the intruder state "works its way down avoided crossings" until in dominates in the description of one of the states in the vacinity of z_d. The same thing is observed for the $|1111\rangle_H$ intruder state for H_4 with EN partitioning. I would prefer if we presented graphs like Fig 8-11 in the H4 paper. **47)Your reason** for having a front door intruder may be confusing. It sounds like we have a front door intruder because of "degeneracy of correlating states", put "$z > 0$" and "front door" close together in the sentence so it is clear what you mean. **48)You write** "In the vicinity of z_d, the eigenfunctions with MP partitioning are dominated by the $|1a_1^2 2a_1^2 1b_2^2\rangle$ (P) and $|1a_1^2 2a_1^2 1b_2 2b_2\rangle$ (Q) CSFs." Please give some numbers here. What are their weights in the wavefunctions. **49)Can this** problem be remove by a better choice of orbitals? These two states differ by a single excitations. A rotation of the two b_2 orbtials can change the importance of the intruder state.

———————————————————————————————————

Figure 3 presents the MP partitioning z-dependent eigenvalue difference for the second and third state of $^1A_1$ symmetry for complex z, i.e., $E_{2^1A_1}(z) - E_{3^1A_1}(z)$. Figure 3 establishes the presence of a degeneracy between these two interspace $^1A_1$ states for z >0, a degeneracy between a pair of states correlating to P and Q space states for z = 0 and therefore corresponding to a "front door" intruder state problem. Figure 3 clearly indicates that the real and imaginary eigenvalues of the second (P space) and third (Q space) states of ${}^1A_1$ symmetry are degenerate for $z_d \approx 0.17 \pm 0.82i$ (since the real and imaginary parts of $[\mathrm{E}_{2^1A_1}(z) - \mathrm{E}_{3^1A_1}(z)] \to 0$ for $z \approx z_d$), leading to a radius of convergence of $R_c = 0.84$ for the perturbation expansion. In the vicinity of z_d, the eigenfunctions with MP partitioning are dominated by the $|1a_1^2 2a_1^2 1b_2^2\rangle$ (P) and $|1a_1^2 2a_1^2 1b_2 2b_2\rangle$ (Q) CSFs. [For $z = 0$, the second excited state (2^1A_1) is, however, described by the CSF $|1a_1^2 2a_1^2 3a_1 4a_1\rangle$]. Hence, choosing a two-state model with this interspace pair of CSFs yields the estimate of $z_d = 0.32 \pm 0.80i$ with $R_c = 0.86$ as the radius of convergence. Thus, we find qualitative agreement between the estimates for the radius of convergence for the full computation and that for the two-state model, although the location of Re z_d differs slightly.

———————————————————————

50)In figure 4 position the labels closer to the curves, so it is clear what curve is what. **51)Use a solid line** for P–corresponding states and dotted for the Q-corresponding ones. Use a smaller font for the labels. See the plots for Be. **52)You** say that in Fig. 4 the avoided crossing is abscured. You must indicate what avoided crossing is obscured. If you claim that one is there we need a way to see it. **53)If you** make an energy difference plot like Fig. 6 in the H4 paper we may be able to see it. Make this plot for Im z = 0. Again, gives the values of the weights of the zeroth order states in the vacinity of z_d. **54)Also** take "(from $z_d \approx -0.06 \pm 0.34i$)" out of parenthesis. It is a key point.

———————————————————————

We now examine the eigenvalue spectrum with barycentric-EN partitioning as a function of the perturbation parameter z, where z is restricted to be real. Figure 4 presents the variation of $[\mathrm{E_m}(z) - \bar{\mathrm{E}}(z)]$ with Re z (for Im $z = 0$). The presence of an avoided crossing for this partitioning is somewhat obscure and is difficult to deduce from the curves for $[\mathrm{E_m}(z) - \bar{\mathrm{E}}(z)]$. However, allowing z to be complex, we again find that the real and imaginary eigenvalues of the second and third ${}^1A_1$ states become degenerate (see Fig. 5) for $z_d \approx -0.06 \pm 0.28i$ ($R_c = 0.29$). Near z_d, the EN partitioning eigenfunctions for these two states are dominated by the $|1a_1^2 2a_1^2 1b_2^2\rangle$ (P space) and $|1a_1^2 2a_1^2 1b_2 2b_2\rangle$ (Q space) CSFs. A two-state model with this interspace pair of CSFs yields the estimated radius of convergence as 0.35 (from $z_d \approx -0.06 \pm 0.34i$), which compares quite favorably with the full 197 CSF calculation.

———————————————————————

55)We should mention beforehand that the single reference computation is divergent. **56)Make it** clear what is the reference space state and at what geometry. **57)Also,** avoided crossings do not necessarily indicate that the degeneracy point is close by. In figure 2 with MP the avoided crossing is close to 1.0. Thus, $|z_d|$ may be greater than one. The same is true for EN in figure 4. We have not shown that $R_c < 1$ for SR perturbation theory.

————————————————————————-

The divergent behavior of the single reference MBPT can also be explained from the complex z-dependent eigenvalue spectra obtained with EN and MP partitioning as follows: Figures 2 (for MP partitioning) and 4 (for EN partitioning) clearly indicate an interspace avoided crossing between the first and second states of $^1A_1$ symmetry. For MP partitioning, the avoided crossing occurs for positive Re z, whereas for EN partitioning it appears at negative Re z. Both yield a radius of convergence $R_c = |z_d| < 1$ and imply a divergent SR perturbation series.

B. Asymptotic convergence with forced valence orbital degeneracy (FD) partitioning when using large reference space

———————————————————————

58)The word "truly" in front of convergent may be too strong. Unless we have searched the complex plane and showed that $R_c > 1$. So tone it down a little. **59)We** have not defined what the larger reference spaces are. This must be added. **60)We** have given the Pade approximants for the larger reference spaces for FD, but we have not given the order-by-order expansions when no Pade is used. Tables should be added for this data, or it is unclear what Pade is doing. **61)It** is important to indicate what orders are needed for each pade approximate. **62)It is written** "We illustrate the behavior only for the three-valence orbital reference space because similar behavior occurs with larger (complete) reference spaces." I do not agree with this, since other problems besides incorrect ordering can cause divergence with FD when valence orbitals are forced degenerate. Zeroth order states can become too close leading to backdoor intruders. Examples are given in the N_2 paper. **63a) We say** that FD is assymptotic. The Pade does not indicate to me that divergence is occuring. This needs to be discussed. **63b) When** does the order-by-order FD results start to deteriorate?

——————————————————————

Tables I through III demonstrate that FD partitioning for the double reference space state (1^1A_1 and 2^1A_1) is truly convergent (through 30th order) over the entire potential surface, while the larger reference space computations yield usefully asymptotic series when truncated at low orders. Tables IV through VI yield well behaved low order Padé approximants for these FD series. The origin of the high order convergence problems for the FD partitioning scheme with large reference spaces can likewise be explained by examining the z- dependent eigenvalues in the complex plane. We illustrate the behavior only for the three-valence orbital reference space because similar behavior occurs with larger (complete) reference spaces.

——————————————————————————-

64)Fig. 6 is very cluttered. Try a different scaling. How about subtracting the eigenvalues by the lowest eivenvalues. Use solid and dotted lines for the P–corresponding and Q-corresponding states, respectively. **65)I do** not even know what the reference space is. What orbitals are valence? **66)Use** the convention that 3^1A_1 is a P–corresponding state—-it is not a P space state! 4^1A_1 is Q corresponding. **67)For** 3^1A_1, put a negative space between the "1" superscript and the "A" and a space after "3" so in looks like $3\,{}^1\!A_1$. **68)You** write "Interestingly, the zeroth order energies of these two interspace states are incorrectly ordered with respect to their expectation values of H in CSF basis, a situation that is expected to produce a "front door" intruder state [see equation (4.6)]." You must first indicate that two zeroth order states, an interspace pair, dominate in the descripption of the exact states that are involved in the avoided crossing in the vicinity of z_{ac}. Then say that these interspace states have incorrect ordering. **69)This** is not surprising. Incorrect ordering divergence is anticipated and has been observer for Be. Please add this point. **70) I would** prefer Fig. 7 to give the energies and not just the energy differences. You finially do mention that two state $|1a_1^2 2a_1^2 1b_1^2\rangle$ (P) and $|1a_1^2 2a_1 3a_1 1b_2^2\rangle$ (Q) states can model the divergent behavior. **71)But we** need some discussion here. These two states are not coupled, so that a two state model actually gives an infinite radius of convergence. However, if the coupling is even infinitesimal than $R_c < 1$. Preferable, we need to consider the 3-state model: Two state incorrectly ordered and not coupled, plus a third state that is coupled to both of them, call these off diagonal matrix elements V1 and V2. Now take the limit of V1 and V2 going to zero. Then $R_c < 1$ for the limit and z_d is real and the crossing appears where the two incorrectly ordered states have their expectation values of H degenerate. **72)You** also write that "Thus, we again find that except in some pathological cases, a simple two-state interspace model provides reasonable information regarding the high REGARDING order convergence behavior of the MR-MBPT procedure and also the states that are responsible for the poor convergence." Please take out the word "pathological". It is jargon that may be unclear to the reader. I prefer just to say that 2–state models can often be used to describe what is going on... Certainly, for large reference spaces, we will find that 2-state models are inadequate. I do not think we should call these larger reference spaces pathological. **73)You** also have "REGARDING" in capitals. There must be some error here. Read over the sentence. However, I do think the points made regarding the use of 2–state models to be worth while. **74)However**, I think we should put this in the discussion and/or introduction.

Figure 6 depicts the variation of $[\mathrm{E_m}(z) - \bar{\mathrm{E}}(z)]$ with Re z (Im $z = 0$) for the first four eigenvalues of $^1A_1$ symmetry using the FD partitioning method and a three orbital reference space. The presence of an avoided crossing is clearly visible in Fig. 6 between the 3^1A_1 (P space) and the 4^1A_1 (Q space) states near Re$z = 0.80$. Interestingly, the zeroth order energies of these two interspace states are incorrectly ordered with respect to their expectation values of H in CSF basis, a situation that is expected to produce a "front door" intruder state [see equation (4.6)]. Detailed analysis confirms that the 3^1A_1 and 4^1A_1 are indeed degenerate (see Fig. 7) for $z_d \approx 0.875 \pm 0.04i$, yielding R_c=0.876. In the vicinity of z_d, these two degenerate states are dominated by the zeroth order $|1a_1^2 2a_1^2 1b_1^2\rangle$ (P) and $|1a_1^2 2a_1 3a_1 1b_2^2\rangle$ (Q) states for FD partitioning. The estimated z_d and R_c for these two interspace state are $\approx$ 0.95 $\pm$ 0.0i and 0.95 which are quite close to those for the full 197 CSFs value. (The imaginary part of z_d is zero because the coupling between these two interspace states vanishes). Thus, we again find that except in some pathological cases, a simple two-state interspace model provides reasonable information regarding the high REGARDING order convergence behavior of the MR-MBPT procedure and also the states that are responsible for the poor convergence. Hence, the interspace two-state models should provide a useful diagnostic for choosing reference spaces in applications of MR-MBPT, just as second order perturbation theory is used for choosing reference configurations for MRCI methods.

————————————————————————————-

75)Why are we performing Pade with the OPT partitioning?

————————————————————————————-

The perturbation expansion are not truly convergent for the FD partitioning scheme with more than two reference configurations, but the series are usefully asymptotic when truncated at low orders and when resummed with [N,M] Padé approximants. We present some of the Padé results in Table V, VI and VII for the OPT partitioning (double reference space) and FD partitioning schemes with three different choices of reference spaces.

VII. RESULTS

————————————————————————————

76)Define MR-MBPT and CCMC. **77)You** say that the FD series are not "truly" convergent. Take out the word "truly", or explain what it means–asymptotic. **78)It** is stated that the FD Pade series are usefully asymptotic. When do these series start to exhibit divergent behavior? **79)It** is observed that the larger reference spaces for FD give better low order results than the smaller ones. I do not agree that this is anticipated. This will not always occur. For example, for N_2 an exhaustive study of the reference space size was undertaken. When the reference space is increased there are many trade-offs that occur. (See the Be paper discussion on FD partitioning.) It can happen that small energy denominators appear when forced degeneracy is imposed. These are generally more severe for larger reference spaces, and, thus, larger reference spaces may give poorer results than the smaller ones. One does not know what is going to happen beforehand. It is very complicted. **80)It** should be mentioned that it takes a very large reference space with FD to match what we get with OPT with 2 valence orbtials. This is very important because a larger reference space is much more computationally expensive.

———————————————————————————

Table VII presents the computed 3rd order $1^1A_1 \rightarrow 2^1A_1$ vertical excitation energy from the forced degeneracy MR-MBPT for various choice of reference spaces, ranging from two to six orbital valence spaces, and from the OPT scheme for the two orbital valence space. We also present the multiconfigurational coupled cluster (CCMC) results of Bannerjee and Simons [57], the quasidegenerate MR-MBPT(QD-MBPT) computation of Lee and Bartlett [54], and the FCI value to evaluate the relative performance of the different approaches. [Since the absolute state energy is not a bound in perturbative methods, Table VII compares excitation energies.]. Tables VIII and IX present the third order $1^1A_1 \rightarrow 3^1A_1$ and $1^1A_1 \rightarrow 4^1A_1$ vertical excitation energies, respectively, as obtained using four, five and six orbital valence spaces. All the FD partitioning computations display the general trend that the accuracy of the third order excitation energies, i.e., the deviation from the FCI excitation energies, improves with an increase in the size of the reference space. For example, the average absolute deviation (for geometries A, B and C) in the third order excitation energy for the $1^1A_1 \rightarrow 2^1A_1$ transition is only 0.024 eV with the six orbital reference space as compared to 0.047 with the five orbital reference space. The improvement is also substantial for higher excited states, even though the larger reference space computations are not truly convergent and diverge more rapidly than smaller ones. These divergent series, however, are transformed to a useful asymptotic form by Padé resummation techniques. Similar patterns are also found for several other geometries along the reaction path of BeH_2. The trends described above are anticipated since an increase in the size of the the reference space improves the first order description and reduces the remaining "correlation contribution" to be obtained perturbatively. Thus, a seven valence orbital reference space produces almost FCI accuracy at third order (but there are then only two excited orbitals in the perturbation expansion) and offers a descent asymptotic perturbation series. Although the perturbative series diverge more rapidly with the larger than the smaller reference spaces (especially for the higher excited states), the high order divergence does not degrade the accuracy of the lower lying states in low orders.

———————————————————————————

81)The paragrah below belongs in the discussion, not the results section. **82)I** prefer we just remove it. **83)I** think most researchers in the field want a single reference space for the entire potential. If we compute gradients to get energy minima, using more that one reference space would not work, or produce serious difficulties, (I think). There will always be a discontinuity in the potential. For dynamics I imagine one needs many points. Then one has to deal with smoothing the surface. It is better to have a single reference space. If we disagree with the "experts" we may be criticized. In this case, I would prefer to avoid the criticism, and pick out "battles" carefully. (Note that Malrieu has written a paper indicating that the FD method is not size extensive.)

———————————————————————————

In order to compute global potential energy surfaces with reasonable accuracy, the use of a large single reference space for all geometries may be extremely difficult and not cost-effective (since the third order computations can become lengthy). While MRCI methods generally use the same reference space for all geometries, this is probably not needed for MR-MBPT method because certain important CSFs in a particular region of geometry may be unimportant in other regions. The Hose-Kaldor incomplete reference space MBPT approach is precisely motivated by the principle of eliminating unimportant CSFs from the reference space. This technique cures the most serious divergence problems of MR-MBPT methods. However, the use of different reference spaces for different regions of geometry is only acceptable if the overall accuracy remains unaltered. (Otherwise the surface might be bumpy.) Moreover, care should be taken to maintain size-extensivity, which is lost with the Hose-Kaldor method. The present FD computations with varying reference spaces shows that the FD method is capable of generating size-extensive global potential energy surface with quite uniform accuracy.

The OPT method converges much faster than with the FD partitioning approach for the two orbital valence space or the QD-MBPT calculations of Bartlett *et al.* [54]. The accuracy of the OPT treatment lies somewhat in between the five and six valence orbital reference space treatments with the FD scheme. Because of this high accuracy, we have not considered OPT treatments with larger reference spaces.

VIII. DISCUSSION

---

108) Define CCMC.

109) We write “This strongly suggests that the accuracy and convergence of the MR-MBPT method depends on the choice of the zeroth order energies, a point noted by many researchers over the past few decades.” I do not think many researchers have realized this. Please remove the the last phrase or change it. I think we are the first to really realize this and exploit it. Paldus has realized it to some extent, but has concluded in his H4 paper that shifting energies can only turn a divergent series into one that converges at a very slow rate. Also, change “This strongly suggests” to “This demonstrates”.

The third order, two-valence orbital space FD calculation closely resemble those from the OPT partitioning method and from the QD-MBPT approach of Lee and Bartlett. These three schemes differ only in the choice of zeroth order energies and the orbitals. All three employ the same number of valence orbitals in the reference space and, hence, the same configuration state functions in the reference space. All three methods generate a truly convergent perturbative series through 30th order. However, the low order (e.g., 3rd order) results differ significantly. For example, the FD partitioning method offers an improved result, i.e., a smaller deviation of the excitation energy from the FCI than from the QD-MBPT method, but an even more significant improvement in the excitation energy is obtained by OPT method. The excitation energy computed with the OPT scheme is comparable to the highly correlated CCMC and FD partitioning method with a six valence orbital reference spaces whose dimension is almost half of the FCI space. Also, the OPT scheme provides a more rapidly convergent perturbative series than the double reference QD-MBPT and FD-MBPT methods. This strongly suggests that the accuracy and convergence of the MR-MBPT method depends on the choice of the zeroth order energies, a point noted by many researchers over the past few decades. [Note that the H_0 formed with unaveraged IVO orbital energies also yields convergent perturbative series and is to some extent better than the FD-series, a finding in accord with small valence space H^v computations demostrating improved results in truly quasidegenerate situations when valence orbital energy averaging may be avoided.]

---

110) The next sentence "An optimal choice of zeroth order energies becomes increasingly difficult $\cdots$" This may only be true if we desire to have a truly convergent series. We have not tried OPT with large reference spaces. Maybe it would work fine.

111) We write "Our FD computations with larger reference spaces suggest that the difficulties in choosing the proper zeroth order energies to enhance the low order convergence can be avoided without sacrificing accuracy (actually we find improved accuracy) by enlarging the reference space in a sensible way." I do not agree with this statement. There is a difficulty with enlarging the reference space: A much larger computation is required. It is always better to use a smaller reference space, if one can get the same accuracy. Again, we have not tried OPT with the larger reference spaces. Also, the main problem with FD and larger reference spaces is the possibility of small energy denominators (see the N2 paper). Thus be removing the constraint of forcing the valence orbitals to be degenerate and choosing the zeroth order energies by some other criteria, say OPT or another way, one can fix up these cases. OPT is the same as FD except we have more degrees of freedom available. It may be more difficult to choose these degrees of freedom for larger reference spaces with OPT, but it is not too difficult to improve on FD, even for the larger reference spaces. (We have written this paper like OPT is developed by someone else and "our method", the method of choice, is FD.)

112) We write "Hence, a large diagonal perturbation matrix element appears due to the forced valence orbital degeneracy restriction, and this perturbation begins contributing third order." Add the other factors as well–problematic orbital energies and incorrect ordering. These points are discussed in detail for the Be paper. Give a citation here. Recall that for N2 when the reference space was increased we often had problems with energy denominators. An important example is the 8OR space where we add the $2\pi_u$ orbitals to the six valence orbitals. It was not possible to remove the energy denominator problems without very large shifts that would make other denominators too large.

113) Can we remove the following: "Of course, situations may arise where the third order correction from the diagonal perturbation is too large to be repaired by arbitrarily increasing the size of the reference space. However, this feature does not necessarily indicate an inefficiency of the method. Therefore, the present computations dispel the pessimistic conclusion [67] that it is not possible to repair the damage to the perturbation expansion caused by the presence of a large diagonal perturbation matrix element due to the forced valence orbital degeneracy." If you prefer to keep it we can discuss it later. I just do not agree that divergence is *entirely* due to a large diagonal V term. For EN we can have a zero diagonal V matrix element, but very poor convergence. (One can show that a type of forced degeneracy and CASPT2 are related, in second order.)

———————————————————————————————————————

An optimal choice of zeroth order energies becomes increasingly difficult and nontrivial to determine as the reference space increases in size. The difficulties are especially severe when the reference space has strong coupling to the virtual space or has incorrectly ordered zeroth order state energies. Our FD computations with larger reference spaces suggest that the difficulties in choosing the proper zeroth order energies to enhance the low order convergence can be avoided without sacrificing accuracy (actually we find improved accuracy) by enlarging the reference space in a sensible way. (Arbitrary choice of the reference space can destroy the convergence and accuracy [65]). Table VIII indicates that as the size of the reference space grows, the accuracy of the FD method increases. The trend in Table VIII is quite expected, since the larger reference space computation offers a better first order description and requires a smaller correlation correction. However, as the size of the reference space increases, the quasidegeneracy of the valence orbital energies diminishes. Hence, a large diagonal perturbation matrix element appears due to the forced valence orbital degeneracy restriction, and this perturbation begins contributing third order. Enlarging the valence space involves a trade-off, and the success of FD method depends on the relative importance of these two opposing factors. When the diagonal perturbation is too large, the convergence and accuracy of FD partitioning may degrade, and an arbitrary increase in the valence space size may not remedy the problem. In fact, the success of the FD scheme lies mostly on the appropriate selection of the reference space, a process which requires some trial and error searching or an a priori knowledge of the most important configurations. Recent computations for methyl mercaptan [50] and hydrogen sulfide [66] potential energy surfaces establish that once the appropriate reference is obtained, the FD method not only provides accurate excitation energies (for both Rydberg and valence states), ionization energies, and dissociation energies but also generates accurate potential energy surfaces even with a limited number of valence space orbitals. Of course, situations may arise where the third order correction from the diagonal perturbation is too large to be repaired by arbitrarily increasing the size of the reference space. However, this feature does not necessarily indicate an inefficiency of the method. Therefore, the present computations dispel the pessimistic conclusion [67] that it is not possible to repair the damage to the perturbation expansion caused by the presence of a large diagonal perturbation matrix element due to the forced valence orbital degeneracy.

IX. CONCLUSIONS

———————————————————————————

114) Change "However, it important to" to "However, it is important to"

115) We write "Further improvements may emerge by optimizing the orbitals by separate unitary transformations ..." I am not aware of this work. I do not think this comment belongs in the conclusion. We have not done any work in this area to make such a conclusion. It belongs in the discussion . I prefer we just remove it, since it is not a critical point. If we keep it we need some citation and examples why this works.

116) We write "as well as the use of more general non-diagonal zeroth order Hamiltonians H_0". I do not think this may be the way to go. CASPT2 uses a diagonal and nondiagonal form of H_0. There are only a few cases where the non diagonal does significantly better than the diagonal. Also, the non diagonal always requires an iterative solution. This makes it much less efficient. There may be a way that a non diagonal can lead to better convergence, but we have not demonstrated here. I think we should remove this point, it especially does not belong in the conclusion section.

117) Please remove the last sentence "The FD approach also appears to be quite useful for generating global potential energy surfaces by varying the set of valence orbital used at different geometries and by piecewise connecting the entire surface without losing accuracy. " We have not done this work in this paper.

———————————————————————————

The present high order multireference perturbative computations further emphasize the utility of two-state models for providing convergence criteria and for assessing the quality of possible choices for partitioning methods, orbitals, etc. An analysis of all interspace two-state models composed of higher lying P-space states and lower lying Q space states, provides an indication of probable impediments to the perturbative convergence and of possible remedies for these problems. This type of analysis bear some relation to the perturbative criterion for selecting reference space configurations for multireference CI methods in its simplicity and general utility. However, it important to mention that situations may arise where the two-state model is inadequate to describe the perturbative divergence problems and where higher dimensional models are required. Situations like this occur when more than one Q space states is important because they are either strongly coupled among themselves or to the higher lying P-space. The optimized partitioning method (OPT) discussed here may not resolve all the problems with MR-MBPT method but may provide a useful alternative tool to tackle the perturbative convergence problem by adjusting the state energies for just the high lying P space and low lying Q space states that enter into the divergent two-state models. Further improvements may emerge by optimizing the orbitals by separate unitary transformations among the highest lying P space and among lowest lying Q space orbitals, as well as the use of more general non-diagonal zeroth order Hamiltonians H_0. Further attempts should be made to optimize the orbital energies instead of the state energies. Optimization of the orbital energies affects more states and, in addition, generates a method that preserves size-extensivity. The forced degeneracy (FD) method is again shown to alleviate the most serious intruder state problems plaguing general multireference methods. The FD approach also appears to be quite useful for generating global potential energy surfaces by varying the set of valence orbital used at different geometries and by piecewise connecting the entire surface without losing accuracy.

X. ACKNOWLEDGMENTS

This research is supported, in part, by the National Science Foundation, Grant No. CH93-07489.

REFERENCES

[1] K. A. Bruckner, Phys. Rev. **97**, 1353 (1955).

[2] J. Goldstone, Proc. R. Soc. London Ser. A **239**, 267 (1957).

[3] N. M. Hugenholtz, Physica **27**, 281 (1957).

[4] J. Hubbard, Proc. R. Soc. London Ser. A **240**, 539 (1957).

[5] F. Cöester, Nucl. Phys. **7**, 421 (1958).

[6] F. Cöester and H. Kümmel, Nucl. Phys. **17**, 477 (1960).

[7] H. Kümmel, K. H. Luhrmann, and J. G. Zabolotzky, Phys. Rep. **36**, 1 (1978).

[8] N. C. Handy, P. J. Knowles, and K. Somasudram, Theor. Chim. Acta **68**, 87 (1985).

[9] R. J. Bartlett and G. D. Purvis, Int. J. Quantum Chem. **14**, 561 (1978).

[10] R. J. Bartlett and G. D. Purvis, Phys. Scr. **21**, 255 (1980).

[11] B. Brandow, Rev. of Mod. Phys. **39**, 771 (1967).

[12] P. Sandars, Adv. Chem. Phys. **14**, 365 (1969).

[13] V. Kvasnička, Czech J. Phys. B **24**, 605 (1974).

[14] I. Lindgren, J. Phys. B **7**, 2441 (1974).

[15] I. Lindgren and J. Morrison, *Atomic Many-Body Theory*, 2nd ed. (Springer, New York, 1986), (Springer Series on Atoms and Plasmas).

[16] G. Oberlechner, F. Owon-N-Cuema, and J. Richert, Nuovo Clemento B **68**, 23 (1970).

[17] T. T. S. Kuo, S. Y. Lee, and K. F. Ratcliff, Nucl. Phys. A **176**, 65 (1971).

[18] M. Johnson and B. Barangar, Ann. Phys. **62**, 172 (1971).

[19] D. M. A. Bannerjee and J. Simons, J. Chem. Phys. **76**, 1979 (1982).

[20] T. H. Schucan and H. A. Weidenmüller, Ann. Phys. N.Y. **73**, 108 (1972).

[21] T. H. Schucan and H. A. Weidenmüller, Ann. Phys. N.Y. **76**, 483 (1973).

[22] G. Hose and U. Kaldor, J. Phys. B **12**, 3827 (1979).

[23] M. Sheppard, J. Chem. Phys. **80**, 1225 (1984).

[24] R. Chaudhuri, D. Sinha, and D. Mukherjee, Chem. Phys. Lett **163**, 165 (1989).

[25] D. Mukherjee, Proc. Ind. Acad. Sci. **96**, 145 (1986).

[26] D. Mukherjee, Chem. Phys. Lett **125**, 207 (1986).

[27] J. P. Malrieu, P. Durand, and J. P. Daudey, J. Phys. A **18**, 809 (1985).

[28] S. Wlison, K. Jankowski, and J. Paldus, Int. J. Quantum. Chem. **23**, 1781 (1983).

[29] S. Wlison, K. Jankowski, and J. Paldus, Int. J. Quantum. Chem. **28**, 525 (1986).

[30] S. Zarrabian and J. Paldus, Int. J. Quantum. Chem. **38**, 761 (1990).

[31] D. M. Silver and R. J. Bartlett, Phys. Rev. A **13**, 1 (1976).

[32] D. M. Silver, S. Wilson, and R. J. Bartlett, Phys. Rev. A **16**, 477 (1977).

[33] C. W. Murray and E. R. Davidson, Int. J. Quantum Chem. **43**, 755 (1992).

[34] K. Andersson, P. Å. Malmqvist, and B. O. Roos, J. Chem. Phys. **96**, 1218 (1992).

[35] M. P. Fülscher, K. Andersson, and B. O. Roos, J. Chem. Phys. **96**, 9204 (1992), and references therein.

[36] K. Wolinski and P. Pulay, J. Chem. Phys **90**, 3647 (1989).

[37] M. R. Hoffman, (Private communication).

[38] D. Mukhopadhyay, B. Datta, and D. Mukherjee, Chem. Phys. Letts **197**, 236 (1992).

[39] J. P. Malrieu, J. L. Heully, and A. Zaitsevski, Theor. Chim. Acta. **90**, 167 (1995).

[40] A. V. Zaitsevskii and J.-L. Heully, J. Phys. B **25**, 603 (1992).

[41] G. Jolicard and G. D. Billing, J. Phys. B **23**, 3457 (1990).

[42] J. P. Heully and J. P. Daudey, J. Chem. Phys. **88**, 1046 (1988).

[43] H. Sun, M. G. Sheppard, and K. F. Freed, J. Chem. Phys. **74**, 6842 (1981).

[44] Y. S. Lee and K. F. Freed, J. Chem. Phys. **77**, 1984 (1982).

[45] X.-C. Wang and K. F. Freed, J. Chem. Phys. **86**, 2899 (1987).

[46] X. C. Wang and K. F. Freed, J. Chem. Phys. **91**, 3002 (1989).

[47] K. F. Freed, *Lecture Notes in Chemistry, Vol.* **52** (Springer, Berlin, 1989), p. 1.

[48] R. L. Graham and K. F. Freed, J. Chem. Phys. **96**, 1304 (1992).

[49] C. H. Martin, R. L. Graham, and K. F. Freed, J. Chem. Phys. **99**, 7833 (1993).

[50] J. E. Stevens, K. F. Freed, M. F. Ardent, and R. L. Graham, J. Chem. Phys. **101**, 4832 (1994).

[51] J. P. Finley and K. F. Freed, J. Chem. Phys. **102**, 1306 (1995).

[52] J. P. Finley, R. K. Chaudhuri, and K. F. Freed, J. Chem. Phys. **103**, 4990 (1995).

[53] J. P. Finley and Rajat K. Chaudhuri and K. F. Freed, Phys. Rev. A (in press).

[54] Y. S. Lee and R. J. Bartlett, Intern. J. Quantum Chem. **S17**, 347 (1983).

[55] G. D. Purvis, R. Sheppard, F. B. Brown, and R. J. Bartlett, Intern. J. Quantum Chem. **23**, 835 (1983).

[56] S. Zarrabian, W. D. Laidig, and R. J. Bartlett, Phys. Rev. A **41**, 4711 (1990).

[57] A. Banerjee and J. Simons, Chem. Phys. **87**, 215 (1984).

[58] M. R. Hoffmann, X.-C. Wang, and K. F. Freed, Chem. Phys. Letts. **136**, 392 (1987).

[59] S. Y. Lee and K. F. Freed, J. Chem. Phys. **104**, 3260 (1996).

[60] R. K. Chaudhuri and K. F. Freed, unpublished work.

[61] B. Huron, P. Rancurel, and J. P. Malrieu, J. Chem. Phys. **58**, 5745 (1973).

[62] S. Wilson, K. Jankowski, and J. Paldus, Int. J. Quantum Chem. **28**, 525 (1985).

[63] J. M. Leinaas and T. T. S. Kuo, Phys. Lett. **62B**, 275 (1976).

[64] P. J. Ellis and E. Osnes, Rev. Mod. Phys. **49**, 777 (1977).

[65] J. E. Stevens, R. K. Chaudhuri, and K. F. Freed, unpublished work on H_2S potential energy surface.

[66] J. E. Stevens and Rajat K. Chaudhuri and K. F. Freed, J. Chem. Phys. (submitted).

[67] Ph. Durand and J. -P. Malrieu, in *Ab initio Methods in Quantum Chemistry*, *Part II*, E. K. Lawley, ed. (Wiley, New York, 1987), p. 321.