Optimization of orbital-specific virtuals in local Møller-Plesset perturbation theory

Yuki Kurashige, Jun Yang, Garnet K.-L. Chan, and Frederick R. Manby

Citation: The Journal of Chemical Physics **136**, 124106 (2012); doi: 10.1063/1.3696962
View online: http://dx.doi.org/10.1063/1.3696962
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/136/12?ver=pdfcov
Published by the AIP Publishing

Articles you may be interested in

An energy decomposition analysis for second-order Møller–Plesset perturbation theory based on absolutely localized molecular orbitals
J. Chem. Phys. **143**, 084124 (2015); 10.1063/1.4929479

Extended Møller-Plesset perturbation theory for dynamical and static correlations
J. Chem. Phys. **141**, 164117 (2014); 10.1063/1.4898804

NMR shielding tensors for density fitted local second-order Møller-Plesset perturbation theory using gauge including atomic orbitals
J. Chem. Phys. **137**, 084107 (2012); 10.1063/1.4744102

Variational second-order Møller–Plesset theory based on the Luttinger–Ward functional
J. Chem. Phys. **120**, 6826 (2004); 10.1063/1.1650307

Convergent summation of Møller–Plesset perturbation theory
J. Chem. Phys. **112**, 4901 (2000); 10.1063/1.481044

Reuse of AIP Publishing content is subject to the terms: https://publishing.aip.org/authors/rights-and-permissions. Downloaded to IP: 147.8.31.43 On: Fri, 02 Sep 2016 04:21:20

Optimization of orbital-specific virtuals in local Møller-Plesset perturbation theory

Yuki Kurashige,[1,a)] Jun Yang,[2] Garnet K.-L. Chan,[2] and Frederick R. Manby[3,b)]
[1]*Department of Theoretical and Computational Molecular Science, Institute for Molecular Science, Okazaki, Aichi, 444-8585, Japan*
[2]*Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York, 14853, USA*
[3]*Centre for Computational Chemistry, School of Chemistry, University of Bristol, Cantocks Close, Bristol, United Kingdom*

(Received 15 November 2011; accepted 6 March 2012; published online 27 March 2012)

We present an orbital-optimized version of our orbital-specific-virtuals second-order Møller-Plesset perturbation theory (OSV-MP2). The OSV model is a local correlation *ansatz* with a small basis of virtual functions for each occupied orbital. It is related to the Pulay–Saebø approach, in which domains of virtual orbitals are drawn from a single set of projected atomic orbitals; but here the virtual functions associated with a particular occupied orbital are specifically tailored to the correlation effects in which that orbital participates. In this study, the shapes of the OSVs are optimized simultaneously with the OSV-MP2 amplitudes by minimizing the Hylleraas functional or approximations to it. It is found that optimized OSVs are considerably more accurate than the OSVs obtained through singular value decomposition of diagonal blocks of MP2 amplitudes, as used in our earlier work. Orbital-optimized OSV-MP2 recovers smooth potential energy surfaces regardless of the number of virtuals. Full optimization is still computationally demanding, but orbital optimization in a diagonal or Kapuy-type MP2 approximation provides an attractive scheme for determining accurate OSVs.
© *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.3696962]

I. INTRODUCTION

In second-order Møller-Plesset perturbation theory (MP2), or in coupled-cluster theory, the number T_{ij}^{ab} represents the probability amplitude for the excitation of a pair of electrons from occupied orbitals i, j to virtual orbitals a, b. Clearly, the number of these rises quartically with the size of the system studied. In canonical electron correlation methods the amplitudes appear to be dense, but there are grounds for optimism that most of the correlation energy can be accounted for using much more compact representations of the information embodied in the amplitudes.

An early approach for compressing correlated-wavefunction expansions was through the use of natural orbitals, that is the basis formed from the eigenvectors of a correlated density matrix.[1,2] The natural orbitals are global to the entire system, and the degree of compression does not take account of the short-range nature of electron correlation. This can be achieved, however, through the use of pair natural orbitals (PNOs), introduced by Edmiston and Krauss[3] and pioneered as a general method for molecular correlation theories in Meyer's development of the coupled-electron-pair approximation.[4] The approach has recently returned to prominence through the work of Neese and co-workers[5,6] and Tew *et al.* have recently demonstrated how the PNO concept can be extended to increase the efficiency of resolution of identity approximations in explicitly correlated F12 theories.[7]

An alternative proposed by Pulay and Saebø is the use of fixed domains of projected atomic orbitals (PAOs) for each (localized) occupied orbital.[8–10] The disadvantages of using a non-orthogonal representation of the virtual space is greatly outweighed by increased locality, and the predetermination of fixed domains has made possible extremely efficient implementations in the groups of Werner and Schütz.[11–15] In particular, linear scaling local implementations of key methods (LMP2,[12,16,17] LCCSD,[15,18] and LCCSD(T) (Refs. 14 and 19)), further enhanced by density fitting (DF-LMP2,[20] DF-LCCSD(T) (Refs. 21 and 22)), have revolutionized the applicability of reliable correlation methods to large molecular systems.[23]

The key criticism of the Pulay–Saebø approach and Werner–Schütz implementations is dependence on a large number of parameters, required to determine the level of theory applied to each occupied-orbital pair, and to determine the domains of PAOs. It has also been noted that without care potentials obtained by these local methods can display discontinuities.[24] It should be pointed out that with attention to detail this can be avoided, and in fact is largely remedied in explicitly correlated variants[25] because the F12 correlation factor can compensate for much of the so-called domain error, as demonstrated by Werner.[26]

In our previous paper,[27] we presented a method in which the level of theory is determined by just a single threshold or integer by means of which the canonical results can be systematically approached, and tested the method on a variety of systems and properties including the dependence on the basis sets and choice of molecular system. It was shown that the orbital-specific virtual (OSV) approximation leads to significant advantages over Pulay–Saebø approaches, both in terms of computational cost and accuracy. In particular, the

a)Electronic mail: kura@ims.ac.jp.
b)Electronic mail: fred.manby@bris.ac.uk.

 © 2012 American Institute of Physics

Reuse of AIP Publishing content is subject to the terms: https://publishing.aip.org/authors/rights-and-permissions. Downloaded to IP: 147.8.31.43 On: Fri, 02 Sep 2016 04:21:20

method produces smooth potential energy curves without reliance on additional system-dependent parameters. In our previous work, we generated sets of OSVs with a very simple procedure based on singular-value decomposition. In this paper, we focus on optimizing the OSVs in order to capture as much of the electron correlation as possible with a given number of virtual orbitals, and to avoid the bottleneck that arises in the SVD procedure for large systems.

II. ORBITAL-SPECIFIC VIRTUAL MP2

In this section, we briefly review the OSV approximation.[27] The method can be summarized by stating that excitations are allowed from a localized molecular orbital to a small number of orbital-specific virtual orbitals. In the language of tensor factorization it can be seen that the method amounts to a decomposition of the full T_2 amplitude in the form,

$$T_{ab}^{ij} \approx T_{\mu_i\nu_j}^{ij} C_{\mu_i a}^{(i)} C_{\nu_j b}^{(j)}, \tag{1}$$

where i, j are occupied orbitals; a, b are virtual orbitals; and μ_i, ν_j are orbital-specific virtuals associated with occupied orbitals i and j, respectively. Here and throughout summation over repeated dummy indices is assumed. Note that Eq. (1) serves only to clarify the relationship between OSV-MP2 and canonical MP2: in OSV-MP2 the canonical amplitudes T_{ab}^{ij} are never explicitly constructed.

This approach is closely related to PAO-local methods, except that there the domains contain functions drawn from a single set of PAOs; here there is the flexibility to consider sets of virtuals $\{\mu_i\}$ specifically suited to the correlation effects in which orbital i participates. The method is also closely related to PNO methods, except here our special virtual orbitals are labelled by a single occupied index, rather than by a pair of occupied orbitals. This can be expected to confer improved efficiency in integral transformation, even when density fitting is used,[5,6] and also makes extension to odd-order excitations more obvious.

In the early days of the PAO-driven local methods the direct excitation approach was used[28] but it was quickly recognized that greater accuracy could be achieved by allowing excitations from orbitals i and j into the *union* of the sets of virtual orbitals associated with each occupied orbital.[10,29] Following these pioneering studies we investigated both the direct (dOSV) and full (fOSV) excitation *ansätze* in our earlier paper.[27]

A. Direct excitation *ansatz*

The first order wavefunction is written as

$$|\Psi^{(1)}\rangle = \frac{1}{2} T_{\mu_i\nu_j}^{ij} \left|\Phi_{\mu_i\nu_j}^{ij}\right\rangle, \tag{2}$$

where

$$\left|\Phi_{\mu_i\nu_j}^{ij}\right\rangle = \hat{E}_{\mu_i i}\hat{E}_{\nu_j j}|\Phi^{(0)}\rangle, \tag{3}$$

where the $\hat{E}_{\mu_i i}$ are spin-summed excitation operators. The Hylleraas MP2 energy functional is given by

$$J = T_{\mu_i\nu_j}^{ij}\left(\tilde{R}_{\mu_i\nu_j}^{ij} + \tilde{K}_{\mu_i\nu_j}^{ij}\right), \tag{4}$$

or

$$J = \tilde{T}_{\mu_i\nu_j}^{ij}\left(R_{\mu_i\nu_j}^{ij} + K_{\mu_i\nu_j}^{ij}\right), \tag{5}$$

where

$$R_{\mu_i\nu_j}^{ij} = K_{\mu_i\nu_j}^{ij} + \left(1 + \hat{P}_{j\nu_j}^{i\mu_i}\right) \times\left[T_{\mu_i\sigma_j}^{ij} f_{\sigma_j\nu_j} - T_{\mu_i\tau_k}^{ik} f^{kj} S_{\tau_k\nu_j}\right], \tag{6}$$

$$K_{\mu_i\nu_j}^{ij} = \langle ij|\mu_i\nu_j\rangle, \tag{7}$$

$$f_{\mu_i\nu_j} = \langle\mu_i|\hat{F}|\nu_j\rangle, \tag{8}$$

$$S_{\mu_i\nu_j} = \langle\mu_i|\nu_j\rangle, \tag{9}$$

and where

$$\tilde{A}_{\mu_i\nu_j}^{ij} = 2A_{\mu_i\nu_j}^{ij} - A_{\nu_j\mu_i}^{ij}. \tag{10}$$

The permutation operator used in Eq. (6) is defined by

$$\hat{P}_{j\nu_j}^{i\mu_i} A_{\mu_i\nu_j}^{ij} = A_{\mu_i\nu_j}^{ij} + A_{\nu_j\mu_i}^{ji}. \tag{11}$$

One can regard the OSV approach as partitioning the virtual space into a space spanned by OSVs $\{\mu_i\}$ and the complementary space spanned by secondary virtual orbitals $\{\bar{\mu}_i\}$ for each occupied orbital i. The amplitudes for excitations to the secondary virtual orbitals are omitted as

$$T_{\bar{\mu}_i\nu_j}^{ij} = T_{\mu_i\bar{\nu}_j}^{ij} = T_{\bar{\mu}_i\bar{\nu}_j}^{ij} = 0. \tag{12}$$

We note that the contravariant amplitudes $\tilde{T}_{\bar{\mu}_i\nu_j}^{ij}$, $\tilde{T}_{\mu_i\bar{\nu}_j}^{ij}$, and $\tilde{T}_{\bar{\mu}_i\bar{\nu}_j}^{ij}$ do not generally vanish under the condition of Eq. (12) because, for example, $\tilde{T}_{\bar{\mu}_i\bar{\nu}_j}^{ij} = 2T_{\bar{\mu}_i\bar{\nu}_j}^{ij} - T_{\bar{\nu}_j\bar{\mu}_i}^{ij}$, and $T_{\bar{\nu}_j\bar{\mu}_i}^{ij}$ is not amongst the set of quantities set to zero.

The same is true for the residual, so the residual equations $\tilde{R}_{\mu_i\nu_j}^{ij} = 0$ and $R_{\mu_i\nu_j}^{ij} = 0$ give different solutions. The latter condition is derived by projection with a set of bra states and does not give an upper bound to the MP2 energy because the contribution of $R_{\nu_j\mu_i}^{ij}$ in Eq. (4) is neglected, and these terms are not constrained to vanish. In practice this makes little difference, but since in this paper we are aiming to optimize the orbital rotations which define the OSVs is becomes important to use the former condition, providing a rigorous upper bound to the MP2 energy.

B. Full excitation *ansatz*

In the full excitation *ansatz* additional cross excitations $ij \rightarrow \nu_j\mu_i$ are included. The first order wavefunction is written as

$$|\Psi^{(1)}\rangle = \frac{1}{2}\sum_{p,q\in\{\mu_i,\nu_j\}} T_{pq}^{ij}\left|\Phi_{pq}^{ij}\right\rangle = \frac{1}{2}\,\mathrm{tr}\begin{pmatrix} T_{\mu_i\sigma_i}^{ij} & T_{\mu_i\tau_j}^{ij} \\ T_{\nu_j\sigma_i}^{ij} & T_{\nu_j\tau_j}^{ij}\end{pmatrix}\begin{pmatrix} \left|\Phi_{\mu_i\sigma_i}^{ij}\right\rangle & \left|\Phi_{\mu_i\tau_j}^{ij}\right\rangle \\ \left|\Phi_{\nu_j\sigma_i}^{ij}\right\rangle & \left|\Phi_{\nu_j\tau_j}^{ij}\right\rangle\end{pmatrix}^T,$$

Reuse of AIP Publishing content is subject to the terms: https://publishing.aip.org/authors/rights-and-permissions. Downloaded to IP: 147.8.31.43 On: Fri, 02 Sep 2016 04:21:20

Thus, the Hylleraas MP2 energy functional is rewritten as

$$J = \mathrm{tr}\begin{pmatrix} \tilde{T}^{ij}_{\mu_i\sigma_i} & \tilde{T}^{ij}_{\mu_i\tau_j} \\ \tilde{T}^{ij}_{\nu_j\sigma_i} & \tilde{T}^{ij}_{\nu_j\tau_j} \end{pmatrix} \times \left[\begin{pmatrix} R^{ij}_{\mu_i\sigma_i} & R^{ij}_{\mu_i\tau_j} \\ R^{ij}_{\nu_j\sigma_i} & R^{ij}_{\nu_j\tau_j} \end{pmatrix} + \begin{pmatrix} K^{ij}_{\mu_i\sigma_i} & K^{ij}_{\mu_i\tau_j} \\ K^{ij}_{\nu_j\sigma_i} & K^{ij}_{\nu_j\tau_j} \end{pmatrix} \right]^T . \quad (13)$$

In contrast to the direct excitation *ansatz*, here each electron in pair ij is excited to the same virtual orbital space, namely, the union $\{\mu_i\} \cup \{\nu_j\}$, so minimization of the Hylleraas functional and the projected equation give the same variational solution.

Because μ_i and ν_j are not always orthogonal, linear dependencies have to be removed. Following the approach adopted in PAO methods,[8–10] we eliminate the redundant vectors by canonical orthogonalization of the overlap matrix S, and truncate small eigenvalues using a threshold of 10^{-6}. Although the active virtual-orbital space is doubled in size by addition of the cross excitation, it is comparable in accuracy to the case where the number of active virtual orbitals is doubled in the direct excitation *ansatz*. The detailed comparison will be shown in Sec. V.

III. ORBITAL OPTIMIZATION IN OSV-MP2

Optimized OSVs are expected to capture as much of the electron correlation as possible with a given number of virtual orbitals. In our previous paper,[27] we generated sets of OSVs by means of singular-value decomposition (SVD) of the diagonal amplitudes

$$\sum_{\mu_i} C^{(i)}_{\mu_i a} \lambda_{\mu_i} C^{(i)}_{\mu_i b} = T^{ii}_{ab} = \frac{K^{ii}_{ab}}{2\epsilon_i - \epsilon_a - \epsilon_b}, \quad (14)$$

where ϵ_i, ϵ_a are orbital energies. When localized orbitals are used for occupied or virtual orbitals, the orbital energies are replaced by diagonal elements of the Fock matrix. Although this provides a reasonable guess, SVD does not necessarily yield optimal OSVs, and in addition, becomes a bottleneck for large systems as will be shown in Sec. IV. The use of optimal OSVs can also be expected to stabilize the numerics, and this is borne out by the calculations presented below.

In this paper, we optimize the shapes of OSVs, i.e., their orbital coefficients, by minimizing the Hylleraas functional in Eq. (4) or (5) with respect to both amplitudes and orbital coefficients of OSVs. The rotations mixing the active and secondary virtual orbitals $\{\psi_p | p \in \mu_i, \bar{\mu}_i\}$ are parametrized as

$$|\psi_{\sigma_i}\rangle = \sum_{p\in\{\mu_i,\bar{\mu}_i\}} |\psi_p\rangle U_{p\sigma_i}, \quad (15)$$

where $\mathbf{U}(\mathbf{X}) = \exp(\mathbf{X})$ and where $\mathbf{X}$ is anti-hermitian: $\mathbf{X} = -\mathbf{X}^\dagger$. Because the Hylleraas functional is invariant to rotations within the active or within the secondary virtual space, the rotation parameters of diagonal blocks $X_{\mu_i\sigma_i}$ and $X_{\bar{\mu}_i\bar{\sigma}_i}$ are redundant and set to be 0. To optimize the orbital coefficients, we use the augmented hessian method[30] in which the Hessian at the reference point $\mathbf{X} = 0$ consists of four blocks

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_{XX} & \mathbf{H}_{XT} \\ \mathbf{H}_{XT}{}^{\mathrm{T}} & \mathbf{H}_{TT} \end{pmatrix}, \quad (16)$$

where

$$H_{X\bar{\mu}_i\sigma_i, X\bar{\nu}_j\tau_j} = \frac{\partial^2 J}{\partial X^{(i)}_{\bar{\mu}_i\sigma_i} \partial X^{(j)}_{\bar{\nu}_j\tau_j}}, \quad (17)$$

$$H_{X\bar{\mu}_i\sigma_i, T^{kl}_{rs}} = \frac{\partial^2 J}{\partial X^{(i)}_{\bar{\mu}_i\sigma_i} \partial T^{kl}_{rs}}, \quad (18)$$

$$H_{T^{ij}_{pq}, T^{kl}_{rs}} = \frac{\partial^2 J}{\partial T^{ij}_{pq} \partial T^{kl}_{rs}}, \quad (19)$$

and where $p \in \{\mu_i\}$, $q \in \{\nu_j\}$, $r \in \{\sigma_k\}$, and $s \in \{\tau_l\}$ in the case of the dOSV *ansatz*, and p, $q \in \{\mu_i, \nu_j\}$ and r, $s \in \{\sigma_k, \tau_l\}$ in the case of the fOSV *ansatz*.

From a technical point of view, the situation is closely analogous to the optimization problem in multiconfigurational self-consistent field theory, but here the number of orbital rotation parameters greatly outnumber the linear parameters because there are many sets of virtual orbitals to be optimized. The H_{TT} block is diagonal dominant and needs only the diagonal elements $H_{T^{ij}_{pq}, T^{ij}_{pq}}$ as a preconditioner. The OSVs are quasi-canonicalized by diagonalizing the diagonal block ($i = j$) of the Fock matrix [Eq. (8)]. Nevertheless, some of the orbital rotation parameters are strongly coupled to each other and some non-diagonal elements of the H_{XX} block, such as $H_{X\bar{\mu}_i\sigma_i, X\bar{\mu}_i\tau_i}$ are necessary for preconditioning. It should be noted that the orbital optimization in the full excitation *ansatz* suffers from intrinsic redundancies between orbital rotations and amplitudes. To remove the redundancies, we have fully diagonalized the hessian matrix and projected out small eigenvalues. Thus, the current implementation of the orbital optimization in the full excitation *ansatz* is impractical and limited to very small molecules, unless these redundancies are removed by using approximate Hylleraas functionals as described in the next paragraph.

In addition to full optimization, we have implemented two different simplified optimization schemes using approximate Hylleraas functionals. The first is based on the Kapuy MP2 (KMP2) approximation,[31–33] which neglects non-diagonal elements of the occupied block of the Fock matrix (or rather, excludes them from the definition of the zeroth-order operator) to give

$$J^{\mathrm{KMP2}} = T^{ij}_{\mu_i\nu_j} \left(\tilde{R}^{\mathrm{KMP2}\,ij}_{\mu_i\nu_j} + \tilde{K}^{ij}_{\mu_i\nu_j} \right), \quad (20)$$

where

$$\tilde{R}^{\mathrm{KMP2}\,ij}_{\mu_i\nu_j} = 2R^{\mathrm{KMP2}\,ij}_{\mu_i\nu_j} - R^{\mathrm{KMP2}\,ij}_{\nu_j\mu_i}, \quad (21)$$

$$R^{\mathrm{KMP2}\,ij}_{\mu_i\nu_j} = K^{ij}_{\mu_i\nu_j} + \left(1 + \hat{P}^{i\mu_i}_{j\nu_j}\right) \times \left[T^{ij}_{\mu_i\sigma_j} f_{\sigma_j\nu_j} - T^{ij}_{\mu_i\nu_j} f^{jj} \right], \quad (22)$$

and

$$R^{\mathrm{KMP2}\,ij}_{\nu_j\mu_i} = K^{ij}_{\nu_j\mu_i} + \left(1 + \hat{P}^{i\mu_i}_{j\nu_j}\right) \times \left[S_{\nu_j\rho_i} \left(T^{ij}_{\rho_i\sigma_j} f_{\sigma_j\mu_i} - T^{ij}_{\rho_i\sigma_j} f^{jj} S_{\sigma_j\mu_i} \right) \right]. \quad (23)$$

Reuse of AIP Publishing content is subject to the terms: https://publishing.aip.org/authors/rights-and-permissions. Downloaded to IP: 147.8.31.43 On: Fri, 02 Sep 2016 04:21:20

TABLE I. CPU time (min) for obtaining OSVs and solving dOSV-MP2 amplitude equations. The number of OSVs is 70 and constant for all calculations, and the error in the correlation energy is always less than around 1%.

		OSV definition		Solve
	N_{basis}	SVD	opt-J^{dMP2}	dOSV-MP2
$(\text{gly})_7$	1192	16	11	34
$(\text{gly})_8$	1354	27	16	48
$(\text{gly})_9$	1516	43	21	67
$(\text{gly})_{10}$	1678	65	29	92
$(\text{gly})_{11}$	1840	94	37	117
$(\text{gly})_{12}$	2002	140	46	153
$(\text{gly})_{13}$	2164	190	57	192
$(\text{gly})_{14}$	2326	245	74	232

The KMP2-optimization expected to be an alternative to the full-optimization, i.e., it is expected that nearly optimal orbitals are obtained. The other is the diagonal MP2 (dMP2) approximation, which neglects some of the non-diagonal elements of the amplitudes ($T^{ij}_{\mu_i\nu_j} = 0$ when $i \neq j$):

$$J^{\text{dMP2}} = T^{ii}_{\mu_i\mu'_i}\left(\tilde{R}^{\text{dMP2}\,ii}_{\mu_i\mu'_i} + \tilde{K}^{ii}_{\mu_i\mu'_i}\right), \tag{24}$$

where

$$R^{\text{dMP2}\,ii}_{\mu_i\mu'_i} = K^{ii}_{\mu_i\mu'_i} + \left(1 + \hat{P}^{\mu_i}_{\mu'_i}\right) \times \left[T^{ii}_{\mu_i\rho_i} F_{\rho_i\mu'_i} - T^{ii}_{\mu_i\mu'_i} f^{ii}\right].$$

In this case, the direct-excitation *ansatz* and full-excitation *ansatz* become equivalent. The dMP2-optimization expected to be an alternative to the SVD procedure.

IV. ACCELERATION OF OSV-MP2

One of the bottlenecks of the previous implementation for large systems is the SVD of the diagonal blocks of the reference amplitudes in Eq. (14), e.g., in Table I, the CPU time for the SVD (245 min) exceeds that for the OSV-MP2 solver (232 min) at a quasi-one-dimensional glycine tetradecamer $(\text{gly})_{14}$.

While the scaling of $O(N_{\text{occ}}N^3_{\text{vir}})$ for obtaining the SVD orbitals of canonical virtual orbitals cannot be reduced by screening due to the delocalized nature of canonical orbitals, this bottleneck can be removed by use of PAOs in Eq. (14), and substituting the eigenvalue ϵ_a with the diagonal Fock-matrix element f_{aa}. Then the K^{ii}_{ab} and (approximate) T^{ii}_{ab} are sparse and the number of relevant PAOs per localized orbital will be constant for large molecules. It should be noted that less correlation energy is captured by these PAO SVD orbitals than canonical SVD orbitals, but the difference can be recovered by optimizing the orbitals in the dMP2 scheme, as shown below. Therefore, the best strategy is to start from PAO SVD orbitals then optimize using the dMP2 approximation; the scaling and computational cost of dMP2 is considerably lower than canonical SVD, as shown in Table I.

Preconditioning for the amplitude equation has been improved. In our previous study, we had simultaneously diagonalized the overlap and Fock matrix in the space of orbital-specific virtuals for each diagonal occupied pair for preconditioning, which is similar to the idea of quasi-canonical orbitals,[11] but still leads to relatively slow convergence for the case of direct excitation *ansatz*, e.g., it takes 16–20 iterations to converge within a tolerance of 10^{-6} hartree. To further accelerate the convergence, diagonal blocks of the Hessian $H_{T^{ij}_{\mu_i\nu_j}, T^{ij}_{\rho_i\sigma_j}}$ are used as a preconditioner, i.e., the linear equation

$$H_{T^{ij}_{\mu_i\nu_j}, T^{ij}_{\rho_i\sigma_j}} \Delta T^{ij}_{\rho_i\sigma_j} = R^{ij}_{\mu_i\nu_j} \tag{25}$$

is solved in each iteration. It is solved by an efficient iterative scheme, dramatically accelerating the convergence of dOSV-MP2. The energy now typically converges to within 10^{-6} hartree in fewer than 8 iterations, comparable to the convergence of PAO-LMP2 with preconditioning using quasi-canonical orbitals.

In many cases, integral transformation is the most time consuming step in LMP2. The density fitting approximation has proven extremely effective for removing this bottleneck, and a linear scaling transformation algorithm can be achieved in PAO-LMP2 by exploiting sparsity of the 3-index integrals, discarding spatially distant occupied pairs, and using the local fitting approximation.[20] The same holds for OSV-MP2, but our current implementation exploits none of these local approximations, so the bottleneck is still the $O(N^5)$ assembly, as for canonical DF-MP2. It should be noted here, unlike the PAO-LMP2, in which active-virtual orbitals are drawn from a single set of PAOs, the OSV-MP2 utilizes as many sets of active-virtual orbitals as the number of occupied orbitals, as its name suggests, and the number of OSVs can be large. The integral transformation

$$(i\nu_j|A) = \sum_p C^{(j)}_{p\nu_j} (ip|A) \tag{26}$$

should be performed for all ν_j for which j forms an occupied pair with i, and is therefore expensive for large systems. The situation is much the same as PNO correlation approaches, and some efficient truncations are required to avoid this bottleneck. We may add in passing that local fitting approximations[20] have been implemented and will be described in a forthcoming paper.

V. RESULTS

To assess the effect of orbital optimization, we performed benchmark calculations for polyglycine and propadienone. The geometries of the polyglycine molecules that are used in this section were generated so as to extend in one-dimensional chains. All calculations have been done on a single Intel Core2Duo 3.0 GHz processor.

A. Accuracy in the correlation energies

Correlation energies of the PAO-LMP2 and OSV-MP2 give upper bounds to the true MP2 correlation energy because the energies of these are obtained by the Hylleraas functional. Therefore, the error in correlation energy is a clear measure of accuracy, and the energy monotonically decreases with the number of virtuals.

Figure 1 gives the error in correlation energy for glycine monomer with cc-pVDZ (Ref. 34) basis sets against the

Reuse of AIP Publishing content is subject to the terms: https://publishing.aip.org/authors/rights-and-permissions. Downloaded to IP: 147.8.31.43 On: Fri, 02 Sep 2016 04:21:20

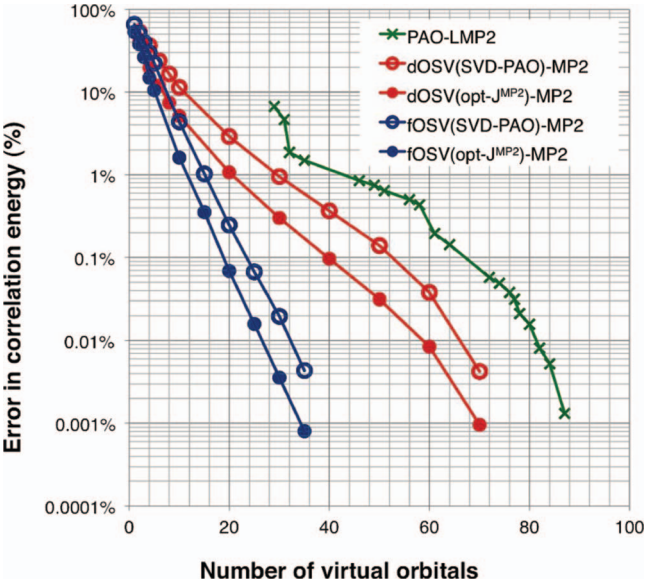

FIG. 1. Errors in the MP2 correlation energy ($-0.800109\ E_h$) against the number of virtual orbitals for glycine monomer with cc-pVDZ basis sets.

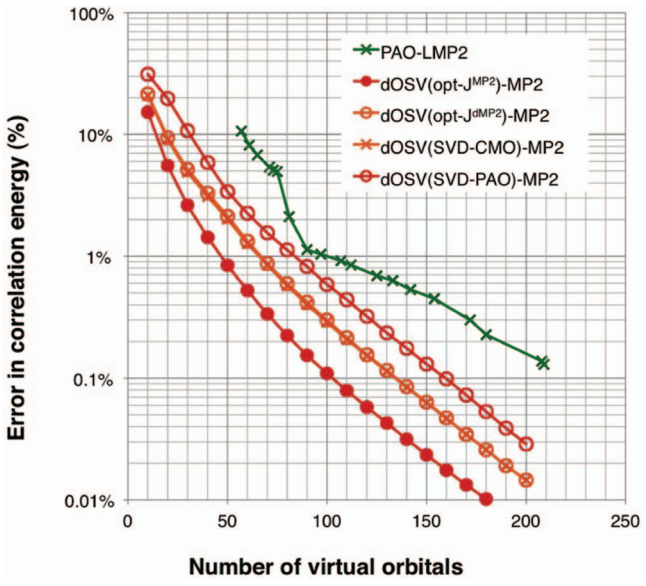

FIG. 2. Errors in the MP2 correlation energy ($-1.756307\ E_h$) against the number of virtual orbitals for glycine dimer with cc-pVTZ basis sets.

number of virtuals, i.e., the average pair domain sizes for PAO-LMP2 and the number of OSVs per occupied orbital for dOSV-MP2 and fOSV-MP2. Note that for fOSV-MP2 the actual active virtual-orbital space per occupied pair is doubled in size by inclusion of the cross excitations; therefore, fOSV-MP2 has roughly twice the computational cost of dOSV-MP2 and PAO-LMP2 calculations with the same number of virtual orbitals in Fig. 1. OSV(SVD-PAO) denotes OSVs obtained by singular value decomposition of diagonal blocks of reference amplitudes using PAOs, as shown in Eq. (14). OSV(opt-J^{MP2}) denotes that OSVs are optimized using the true Hylleraas functional.

It can be seen that OSV-MP2 is consistently more accurate than PAO-LMP2 even with the simple, non-optimal SVD orbitals. With both the direct and full excitation *ansätze*, the accuracy of the OSV-MP2 was much improved by optimizing OSVs; in other words, the number of virtuals necessary to get a given accuracy is reduced by optimization. This small system is nearly the maximum limit for the current implementation of orbital optimization with the fOSV-MP2 *ansatz* because there are remaining redundancies as described above, and an efficient algorithm to optimize simultaneously the orbitals and amplitudes has not yet been found.

Figure 2 gives the error in correlation energy for a larger system, glycine dimer, using the cc-pVTZ (Ref. 34) basis set. OSV(SVD-CMO) denotes OSVs obtained by singular value decomposition of diagonal blocks of reference amplitudes using canonical virtual orbitals [See Eq. (14)]. OSV(opt-J^{dMP2}) denotes that the OSVs are optimized by minimizing the approximate dMP2 Hylleraas functional of Eq. (24). The OSV(SVD-CMO) and OSV(SVD-PAO) are not identical because in Eq. (14) the eigenvalues are simply replaced by diagonal Fock-matrix elements for PAOs. OSV(SVD-CMO)-MP2 consistently gave lower energies than OSV(SVD-PAO)-MP2 for a given number of virtual orbitals, as shown in Fig. 2.

Interestingly, OSV(opt-J^{dMP2}) gives more-or-less the same energies as OSV(SVD-CMO), but orbital optimization based on the dMP2 *ansatz* is more efficient than the expensive singular value decomposition, which scales as $O(N_{\mathrm{occ}}N_{\mathrm{vir}}^3)$. Naturally, the fully optimized OSVs further improve these results, albeit at considerable computational cost.

Figure 3 and Table II give the error in correlation energy and the CPU time for the orbital optimization, respectively, for larger polyglycine molecules up to the hexamer with the cc-pVTZ basis. OSV(opt-J^{KMP2}) denotes the OSV that is optimized by using the approximate Hylleraas functional in Eq. (20).

The number of OSVs was 60 for all calculations. Figure 3 shows that errors remained below 1% for

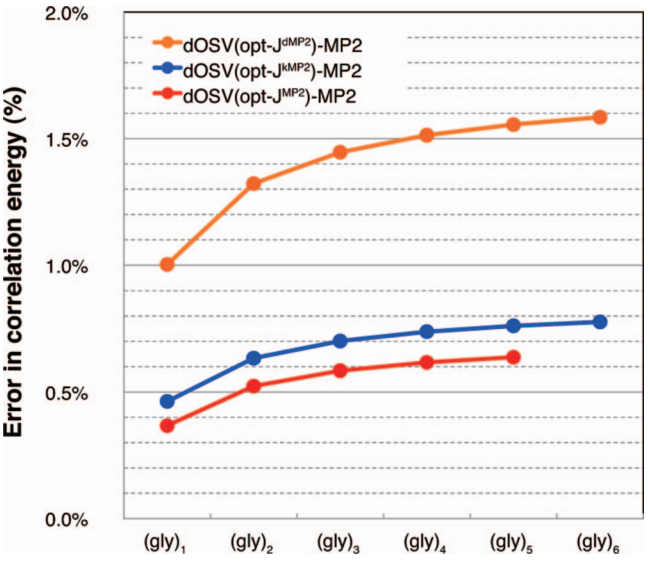

FIG. 3. Error in correlation energy with OSV(opt-J^{dMP2}), OSV(opt-J^{KMP2}), and OSV(opt-J^{MP2}) for $(\mathrm{gly})_{n=1-6}$. The number of active virtuals ψ_{μ_i} is 60 and constant for all the calculations.

Reuse of AIP Publishing content is subject to the terms: https://publishing.aip.org/authors/rights-and-permissions. Downloaded to IP: 147.8.31.43 On: Fri, 02 Sep 2016 04:21:20

TABLE II. CPU time (s) for optimizing OSVs averaged over all macro iterations (orbital rotations).

	$(gly)_1$	$(gly)_2$	$(gly)_3$	$(gly)_4$	$(gly)_5$	$(gly)_6$
N_{basis}	220	382	544	706	868	1030
Integral transformation						
	6	56	240	649	1472	3599
Augmented Hessian						
opt-J^{KMP2}	91	355	749	1274	1883	3547
opt-J^{MP2}	442	2779	6595	11 159	23 359	...

OSV(opt-J^{KMP2}) and OSV(opt-J^{MP2}), and below 2% for OSV(opt-J^{dMP2}). The errors rapidly saturate for large molecules due to the spatial locality of electron correlation. The error in correlation energy changed by only 0.01% between the pentamer and hexamer for OSV(opt-J^{KMP2}), and by 0.03% for OSV(opt-J^{dMP2}). This indicates that the errors of OSV(opt-J^{MP2}) and OSV(opt-J^{KMP2}) are saturated somewhat more rapidly than those of OSV(opt-J^{dMP2}).

As shown in Table II, the CPU time for orbital optimization is greatly reduced by the use of the KMP2 approximation, which neglects non-diagonal elements of the occupied block of the Fock matrix in the Hylleraas functional. The errors for OSV(opt-J^{KMP2}) are close and almost parallel to those for OSV(opt-J^{MP2}). The CPU time for the augmented-hessian eigenvalue equation solver of OSV(opt-J^{KMP2}) scales with system size as $O(N^{2.2})$ while that of OSV(opt-J^{MP2}) scales as $O(N^{2.8})$. This is because the most expensive step in OSV(opt-J^{MP2}) of $O(N_{occ}^3 N_{OSV}^3)$ is reduced to $O(N_{occ}^2 N_{OSV}^3)$ by the KMP2 approximation. Therefore, OSV(opt-J^{KMP2}) is an efficient alternative to OSV(opt-J^{MP2}), or can be used to obtain a close initial guess for the more accurate method. The integral transformation step scales steeply as $O(N^{4.0})$ and becomes a bottleneck in OSV(opt-J^{KMP2}) for large systems, although as mentioned earlier, this bottleneck can be essentially removed through exploiting integral sparsity.

B. Potential energy curves

While the error in correlation energy is a good measure of accuracy of the approximations, relative energies of different geometries or states are obviously of greater practical importance.

Figure 4 shows potential energy curves for dissociation of the $CH_2C{=}CO$ molecule by the dOSV-MP2 and fOSV-MP2 methods using various types of OSVs. Each geometry was optimized at the MP2/cc-pVDZ level with frozen C=C bond length. The energies relative to the dissociation limit are shown in kcal/mol. The true MP2 curve was nicely reproduced by the dOSV-MP2 method, using fewer than half the number of virtual orbitals ($N_{OSV} = 25$ cf. $N_{vir} = 52$), with errors in dissociation energy of 1.4 and 4.0 kcal/mol using OSV(opt-J^{MP2}) and OSV(SVD-CMO) virtuals, respectively.

In Ref. 24 it was shown that PAO-LMP2 exhibits a discontinuous potential energy curve for this system. Although the discontinuity problem seemed to be fixed by OSV-MP2 in our previous study,[27] it reappears for dOSV-MP2(SVD-CMO) and fOSV-MP2(SVD-CMO) when the extremely small number of active virtual-orbitals were used, i.e., the calculations with $N_{OSV} = 2$. These energy jumps occur when SVD eigenvalues cross and where only one of the two degenerate eigenfunctions are included in the OSV space. However, the potential energy curves of OSV(opt-J^{MP2})-MP2 are continuous by definition, and even with this small number of virtuals was surprisingly parallel to the true MP2 curve. The most stable bond lengths at HF, MP2, and OSV(opt-J^{MP2})-MP2 with $N_{OSV} = 2$ were 1.29, 1.34, and 1.33Å, respectively.

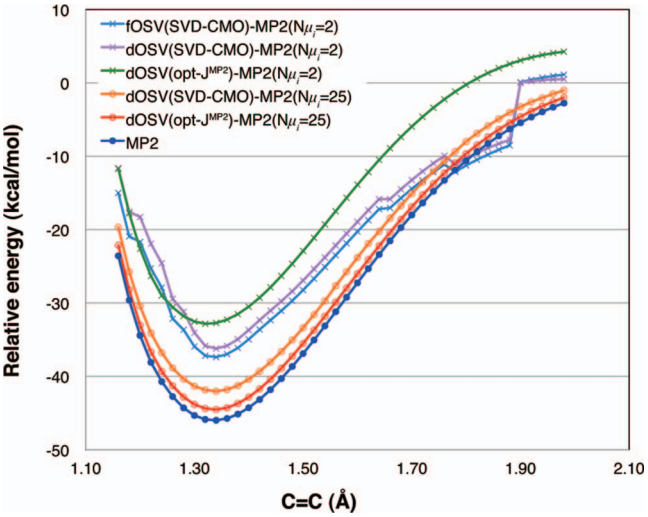

FIG. 4. Potential energy curves resulting from various OSV-MP2 approximations for extension of the central C=C bond of propadienone using a cc-pVDZ basis set. Each geometry was optimized at the MP2/cc-pVDZ (52 virtual orbitals) level with frozen C=C bond length, and relative energies of the total energies from the dissociation limits are shown in kcal/mol.

VI. SUMMARY

Orbital-specific virtual MP2 has been improved by optimizing the virtual orbitals. In our previous study, OSVs for occupied orbital i were generated by SVD of the diagonal block of reference amplitudes, T_{ab}^{ii}. Here we have simultaneously optimized amplitudes and OSVs by minimizing the MP2 Hylleraas functional or approximations to it that neglect various coupling terms. In addition, we have accelerated the OSV-MP2 solver by implementing the method with density-fitting and with a new preconditioner for dOSV-MP2. The simple SVD step, which is a bottleneck in our previous method, has been replaced by efficient dMP2 optimization. Accuracy in the correlation energies of the polyglycine molecules is considerably improved by use of MP2-optimized orbitals.

ACKNOWLEDGMENTS

This work was begun during a visit of one of us (Y.K.) to Bristol in 2010. Y.K. acknowledges support for this work from Ministry of Education, Culture, Sports, Science and Technology-Japan (MEXT) via the Young Scientists (B) (Grant No. 21750028) G.K.-L.C. acknowledges support for

Reuse of AIP Publishing content is subject to the terms: https://publishing.aip.org/authors/rights-and-permissions. Downloaded to IP: 147.8.31.43 On: Fri, 02 Sep 2016 04:21:20

this work from the U.S. Department of Energy (DOE), Office of Science via Award No. DE-FG02-07ER46432, as well as support from the David and Lucile Packard foundation.

[1]A. C. Hurley, J. Lennard-Jones, and J. A. Pople, Proc. R. Soc. London, Ser. A **220**, 446 (1953).
[2]P.-O. Löwdin and H. Shull, Phys. Rev. **101**, 1730 (1956).
[3]C. Edmiston and M. Krauss, J. Chem. Phys. **42**, 1119 (1965).
[4]W. Meyer, J. Chem. Phys. **58**, 1017 (1973).
[5]F. Neese, A. Hansen, and D. G. Liakos, J. Chem. Phys. **131**, 064103 (2009).
[6]F. Neese, F. Wennmohs, and A. Hansen, J. Chem. Phys. **130**, 114108 (2009).
[7]D. P. Tew, B. Helmich, and C. Hättig, J. Chem. Phys. **135**, 074107 (2011).
[8]P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).
[9]P. Pulay and S. Saebø, Theor. Chim. Acta **69**, 357 (1986).
[10]S. Saebø and P. Pulay, J. Chem. Phys. **86**, 914 (1987).
[11]C. Hampel and H.-J. Werner, J. Chem. Phys. **104**, 6286 (1996).
[12]M. Schütz, G. Hetzer, and H.-J. Werner, J. Chem. Phys. **111**, 5691 (1999).
[13]M. Schütz, Phys. Chem. Chem. Phys. **4**, 3941 (2002).
[14]M. Schütz and H.-J. Werner, Chem. Phys. Lett. **318**, 370 (2000).
[15]M. Schütz and H.-J. Werner, J. Chem. Phys. **114**, 661 (2001).
[16]G. Hetzer, P. Pulay, and H.-J. Werner, Chem. Phys. Lett. **290**, 143 (1998).
[17]G. Hetzer, M. Schütz, H. Stoll, and H.-J. Werner, J. Chem. Phys. **113**, 9443 (2000).
[18]C. Hampel, K. A. Peterson, and H.-J. Werner, Chem. Phys. Lett. **190**, 1 (1992).
[19]M. Schütz, J. Chem. Phys. **113**, 9986 (2000).
[20]H.-J. Werner, F. R. Manby, and P. J. Knowles, J. Chem. Phys. **118**, 8149 (2003).
[21]M. Schütz and F. R. Manby, Phys. Chem. Chem. Phys. **5**, 3349 (2003).
[22]H.-J. Werner and M. Schütz, J. Chem. Phys. **135**, 144116 (2011).
[23]F. Claeyssens, J. N. Harvey, F. R. Manby, R. A. Mata, A. J. Mulholland, K. E. Ranaghan, M. Schütz, S. Thiel, W. Thiel, and H.-J. Werner, Angew. Chem. **118**, 7010 (2006).
[24]N. J. Russ and T. D. Crawford, J. Chem. Phys. **121**, 691 (2004).
[25]H.-J. Werner, T. B. Adler, G. Knizia, and F. R. Manby, in *Recent Progress in Coupled Cluster Methods* (Springer, Dordrecht, 2010), pp. 573–619.
[26]H.-J. Werner, J. Chem. Phys. **129**, 101103 (2008).
[27]J. Yang, Y. Kurashige, F. R. Manby, and G. K.-L. Chan, J. Chem. Phys. **134**, 044123 (2011).
[28]P. Pulay, S. Saebø, and W. Meyer, J. Chem. Phys. **81**, 1901 (1984).
[29]S. Saebø and P. Pulay, Annu. Rev. Phys. Chem. **44**, 213 (1993).
[30]B. H. Lengsfield III, J. Chem. Phys. **73**, 382 (1980).
[31]E. Kapuy, Z. Cs'epes, and C. Kozmutza, Int. J. Quantum Chem. **23**, 981 (1983).
[32]E. Kapuy, F. Bartha, F. Bog'ar, Z. Cs'epes, and C. Kozmutza, Int. J. Quantum Chem. **38**, 139 (1990).
[33]J. E. Subotnik and M. Head-Gordon, J. Chem. Phys. **122**, 034109 (2005).
[34]T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

Reuse of AIP Publishing content is subject to the terms: https://publishing.aip.org/authors/rights-and-permissions. Downloaded to IP: 147.8.31.43 On: Fri, 02 Sep 2016 04:21:20