Local orbitals by minimizing powers of the orbital variance

Branislav Jansík,[a] Stinne Høst, Kasper Kristensen, and Poul Jørgensen
Lundbeck Foundation Center for Theoretical Chemistry, Department of Chemistry, University of Aarhus, DK-8000 Århus C, Denmark

(Received 21 March 2011; accepted 23 April 2011; published online 16 May 2011)

It is demonstrated that a set of local orthonormal Hartree–Fock (HF) molecular orbitals can be obtained for both the occupied and virtual orbital spaces by minimizing *powers* of the orbital variance using the trust-region algorithm. For a power exponent equal to one, the Boys localization function is obtained. For increasing power exponents, the penalty for delocalized orbitals is increased and smaller maximum orbital spreads are encountered. Calculations on superbenzene, C_{60}, and a fragment of the titin protein show that for a power exponent equal to one, delocalized outlier orbitals may be encountered. These disappear when the exponent is larger than one. For a small penalty, the occupied orbitals are more local than the virtual ones. When the penalty is increased, the locality of the occupied and virtual orbitals becomes similar. In fact, when increasing the cardinal number for Dunning's correlation consistent basis sets, it is seen that for larger penalties, the virtual orbitals become *more local* than the occupied ones. We also show that the local virtual HF orbitals are significantly more local than the redundant projected atomic orbitals, which often have been used to span the virtual orbital space in local correlated wave function calculations. Our local molecular orbitals thus appear to be a good candidate for local correlation methods. © *2011 American Institute of Physics*. [doi:10.1063/1.3590361]

I. INTRODUCTION

In a standard Hartree–Fock (HF) calculation, a set of canonical molecular orbitals (MOs) are obtained from diagonalization of the Fock eigenvalue equation. The canonical HF orbitals constitute the backbone for our qualitative understanding of the electronic structure of a molecular system – for example, according to Koopmans' theorem, the ionization energy and the electron affinity of a molecule can be approximated by canonical orbital energies for the highest occupied MO (HOMO) and lowest unoccupied MO (LUMO), respectively.

For a quantitative description, the mean-field HF model is insufficient, and more accurate electronic structure models have to be used to describe the correlated movement of the electrons. Using a simplistic physical description, electron correlation effects may be described in terms of the very short-ranged electron-electron interactions leading to Coulomb holes in the wave function, and the slightly longer ranged interactions leading to dispersion effects. To describe these local phenomena in an efficient manner, a basis of local orbitals has to be used. The canonical HF orbitals are very delocalized, and they are therefore very inefficient to use for describing the electron correlation. In fact, when the calculation of the correlation energy is expressed in terms of the canonical basis, a scaling wall is encountered for larger systems.

The HF energy is invariant under rotation among the occupied and among the unoccupied MOs, and this can be used to transform the canonical orbitals to a local HF basis.[1] The simple qualitative interpretation associated with the canonical orbitals is then lost, but the correlated movement of electrons becomes more easily described and interpreted. In this paper we discuss how a set of local occupied and unoccupied HF orbitals can be obtained by optimizing localization functions.

A large variety of different localization procedures have been proposed.[2–15] Three localization functions are commonly used. The scheme attributed to Boys[2–4] minimizes the spatial extent of the MOs by maximizing the distance between the orbital centroids. This is equivalent to maximizing the function

$$\xi_{\text{Boys}} = \sum_p \sum_x \langle p|\hat{x}|p\rangle^2, \tag{1}$$

where $|p\rangle$ refers to a set of occupied (or unoccupied) MOs and $\hat{x}$ is a component of the position operator. The scheme by Edmiston–Ruedenberg (ER) (Refs. 5–7) maximizes the self-repulsion energy

$$\xi_{\text{ER}} = \sum_p (pp|pp) \tag{2}$$

and the scheme by Pipek–Mezey (PM) (Refs. 11 and 12) minimizes the number of atomic centers over which each MO extends by maximizing the sum of the squares of the gross atomic Mulliken population of the MOs,

$$\xi_{\text{PM}} = \sum_p \sum_A |\langle p|\hat{P}_A|p\rangle|^2, \tag{3}$$

where $\hat{P}_A$ is the projection operator onto the space of atomic orbitals (AOs) centered on atom A. Reference 12 gives a brief introduction to the different localization functions and the Jacobi sweep that is usually used to obtain the localized MOs.

For the occupied molecular orbitals, the above localization functions usually have strong and isolated minima, and

[a]Electronic mail: jansik@chem.au.dk.

 © 2011 American Institute of Physics

Downloaded 15 Jun 2011 to 130.225.22.254. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/about/rights_and_permissions

the localization functions have been successfully optimized using a Jacobi sweep of iterations to give a set of local occupied HF orbitals. *For the virtual space by contrast, optimization of localization functions does not converge rapidly to a minimum. Convergence is slow and painful. There appear to be many local minima and these are not well separated*, as stated by Subotnik *et al.*[16] In general, localized virtual HF orbitals cannot be obtained using the above localization functions and a Jacobi sweep of iterations.

In the absence of localized virtual HF orbitals, projected atomic orbitals (PAOs), where the occupied orbital space is projected out of the AO basis, have been used to span the virtual orbital space.[17–32] Besides being redundant, a more severe drawback of the PAOs is that they are much less local than the localized occupied HF orbitals.

We have recently obtained local occupied and local virtual HF orbitals by applying the least-change approach[33,34] in combination with the three-level optimization algorithm[35,36] for the HF energy. The least-change algorithm utilizes the fact that the solution to the Fock eigenvalue equation can be divided into three parts: (1) a transformation from the AO basis to an orthonormal basis, (2) a Hartree–Fock optimization condition in the orthonormal basis, which ensures that the Fock matrix is block-diagonal with vanishing elements between the occupied and unoccupied HF orbitals, and (3) a canonical condition, which ensures that the Fock matrix is diagonal with orbital energies on the diagonal. In the least-change approach, the canonical condition is removed from the solution of the Fock eigenvalue equation, and the smallest possible transformation matrices are determined from the AO basis to its orthogonalized counterpart and from this orthogonalized basis to an optimized HF basis. The least-change approach gives a set of local orbitals because small transformations are carried out on the local AO basis. The least-change approach may be applied to both the valence basis optimization and the full basis optimization in the three-level optimization algorithm, giving a set of local occupied and virtual HF orbitals.[33] The major problem with the least-change algorithm is that in Part 2, the identification of the smallest transformation matrix depends on the ordering of the orthonormal orbitals, i.e., whether an orthonormal basis function is associated with the occupied or the unoccupied orbital space. The optimal ordering can be identified by examining the orbital spread for all different orderings. The ordering that leads to the smallest maximum orbital spread gives the least-change molecular (LCM) orbitals. When determining the LCM orbitals for larger molecular systems, the number of orderings becomes prohibitively large. Below we describe an algorithm which may be carried out also for large basis sets.

Maximizing the function in Eq. (1) is equivalent to minimizing the sum of the orbital variances,

$$\xi = \sum_p \sum_x (\langle p| (\hat{x} - \langle p|\hat{x}|p\rangle)^2 |p\rangle) = \sum_p \sum_x (\langle p|\hat{x}^2|p\rangle - \langle p|\hat{x}|p\rangle^2), \quad (4)$$

because $\sum_p \langle p|\hat{x}^2|p\rangle$ is invariant to orbital rotations among the occupied and among the unoccupied orbitals. Minimizing Eq. (4) does not generate a set of local virtual HF orbitals when a Jacobi sweep of iterations is used. However, in this paper we demonstrate that using the trust-region minimization algorithm of Fletcher,[37] a set of local virtual HF orbitals can be obtained. As starting orbitals we use a simplified version of the LCM orbitals, where the expensive reordering is replaced by a simpler reordering scheme based on the overlap between the orthonormal orbitals and the set of occupied orbitals. These simplified LCM orbitals are relatively local and therefore provide a much better starting guess than the canonical orbitals.

The localization function in Eq. (4) determines the set of occupied and virtual HF orbitals which on average are most local. However, in a correlated wave function context, the locality of a set of orthonormal MOs is determined by the locality of the *least* local orbitals. To make the least local MOs more local, we may introduce a *penalty* for delocalized orbitals. In practice this may be done by using localization functions, which contain *powers* of the orbital variance as a measure of locality. We investigate what happens when powers of the orbital variance is used in Eq. (4) instead of the simple orbital variance. We also compare the locality of our occupied HF orbitals to Boys orbitals and LCM orbitals, and for our virtual HF orbitals we compare to PAOs and LCM orbitals.

In Sec. II, we describe our locality function, where penalties are introduced for delocalized orbitals. We also briefly summarize the trust-region method, which is used for optimizing the localization function to obtain local occupied and unoccupied HF orbitals. We describe how the level-shifted Newton equations entering the trust-region method may be solved using iterative schemes and derive the linear transformation of the Hessian on a trial vector, which is required when solving the Newton equations. Numerical test examples are given in Sec. III to illustrate the locality of the obtained occupied and virtual HF orbitals. Finally, in Sec. IV we summarize our developments and put them in a broader perspective.

II. LOCAL ORBITALS USING POWERS OF THE ORBITAL VARIANCE

A. Orbital variance as a localization measure

The *orbital variance* Ω_p for an orbital $|p\rangle$ is defined as its average deviation from its average position squared,

$$\Omega_p = \sum_x (\langle p| (\hat{x} - \langle p|\hat{x}|p\rangle)^2 |p\rangle) = \sum_x (\langle p|\hat{x}^2|p\rangle - \langle p|\hat{x}|p\rangle^2), \quad (5)$$

and the corresponding *orbital spread* σ_p is the square root of this number

$$\sigma_p = \Omega_p^{1/2}. \quad (6)$$

The orbital variance (or spread) of orbital $|p\rangle$ is a measure of the spatial extent of that orbital and thus of the locality of the orbital. The sum of the variances for a set of orthonormal

Downloaded 15 Jun 2011 to 130.225.22.254. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/about/rights_and_permissions

orbitals is a measure of the locality for this set

$$\xi = \sum_p \Omega_p. \tag{7}$$

Minimizing ξ while keeping the orbitals orthonormal gives the set of orbitals that on average are most local.

In local correlation electronic structure methods it is important that the *average* orbital variance is small. However, it is even more important that the *maximum* orbital variance is small, because a single delocalized orbital will destroy the locality of the set of orbitals as a whole. A small maximum orbital variance can be obtained by introducing a penalty on the orbitals with large orbital variances. This may be accomplished by using powers of the variance,

$$\xi_m = \sum_p \Omega_p^m, \tag{8}$$

as a measure of locality, where m is a positive integer. Increasing m increases the penalty for having orbitals with a large orbital variance.

In the following we assume that we have a set of orthonormal occupied or virtual HF orbitals and determine the set that minimizes ξ_m. We investigate the effect of increasing the penalty (value of m) on the orbitals having a large variance, noting that in general, a decrease in the variance of one orbital is obtained at the expense of increasing the orbital variance of another orbital. We also note that Eq. (8) for $m = 1$ reproduces Eq. (4), and it is therefore equivalent to the Boys localization function.

In a HF calculation, the molecular core orbitals are in general almost identical to the atomic core orbitals. If the molecular core orbitals are included in the summation in Eq. (8), the molecular valence orbitals will get undesirable small tail coefficients referencing the core orbitals. We therefore exclude the molecular core orbitals from the summation in Eq. (8) to maintain the division of electrons into core and valence spaces.

To minimize ξ_m, we will use an orthogonal parameterization of the orbitals

$$|\tilde{p}\rangle = \exp(-\hat{\kappa}) a_{p\sigma}^\dagger |vac\rangle, \tag{9}$$

$$\hat{\kappa} = \sum_{i>j} \kappa_{ij}(E_{ij} - E_{ji}) = \sum_{ij} \kappa_{ij} E_{ij}, \tag{10}$$

$$E_{ij} = a_{i\alpha}^\dagger a_{j\alpha} + a_{i\beta}^\dagger a_{j\beta}, \tag{11}$$

where $a_{p\sigma}^\dagger$ is a creation operator with σ denoting spin. The summations in Eq. (10) refer either to the set of occupied orbitals (excluding core orbitals) or to the set of virtual orbitals. We thus parameterize ξ_m in Eq. (8) as

$$\tilde{\xi}_m(\kappa) = \sum_p \tilde{\Omega}_p^m, \tag{12}$$

where

$$\tilde{\Omega}_p^m = \left(\sum_x \langle \tilde{p}|\hat{x}^2|\tilde{p}\rangle - \langle \tilde{p}|\hat{x}|\tilde{p}\rangle^2 \right)^m. \tag{13}$$

B. Trust-region minimization

The minimum of $\tilde{\xi}_m$ may be determined using a trust-region optimization. We refer to Ref. 37 for a detailed discussion of the trust-region method. Here, we summarize the important features of the method.

The trust-region method is built up around a second-order Taylor series $Q(\boldsymbol{\kappa})$ of $\tilde{\xi}_m$ in Eq. (12),

$$Q_m(\boldsymbol{\kappa}) = \xi_m^{[0]} + \boldsymbol{\kappa}^T \boldsymbol{\xi}_m^{[1]} + \frac{1}{2}\boldsymbol{\kappa}^T \boldsymbol{\xi}_m^{[2]} \boldsymbol{\kappa}, \tag{14}$$

where $\xi_m^{[0]}$ is the value of $\tilde{\xi}_m$ at the expansion point and $\boldsymbol{\xi}_m^{[1]}$ and $\boldsymbol{\xi}_m^{[2]}$ are the gradient and Hessian of $\tilde{\xi}_m$, respectively, at the expansion point. We assume $\boldsymbol{\kappa} = \mathbf{0}$ at the expansion point. The idea behind the trust-region method is to take steps only in the region where the second-order Taylor expansion of $\tilde{\xi}_m$ can be trusted, i.e., where the second-order expansion is a good representation of the full expansion. This region defines the trust-region of the second-order expansion and is a hypersphere of radius h around the expansion point. The trust-region is updated during the iterative procedure depending on how quadratic the surface is. Let us initially describe how the optimal step is taken in the trust-region method for a trust-radius h and then go on to describe how h is updated.

If the Hessian is positive definite and the step is smaller than the trust radius h, then the Newton step is taken

$$\boldsymbol{\xi}_m^{[2]} \boldsymbol{\kappa} = -\boldsymbol{\xi}_m^{[1]}, \tag{15}$$

as it can be trusted leading to the minimum of $Q(\boldsymbol{\kappa})$. At the end of an iteration sequence, the Newton steps always become acceptable and the iteration sequence will converge quadratically to the minimum of ξ_m.

If the Newton step goes beyond the trust-region, or if the Hessian is not positive definite, then the step giving the minimum value of the second-order Taylor series Q_m in Eq. (14) on the boundary of the trust-region is taken. To determine this step, we determine the minimum of the second-order surface in Eq. (14) subject to the condition that the step has to be equal to the trust-radius

$$\| \boldsymbol{\kappa} \| = h. \tag{16}$$

We therefore set up the Lagrangian

$$L(\boldsymbol{\kappa}, \mu) = \xi_m^{[0]} + \boldsymbol{\kappa}^T \boldsymbol{\xi}_m^{[1]} + \frac{1}{2}\boldsymbol{\kappa}^T \boldsymbol{\xi}_m^{[2]} \boldsymbol{\kappa} - \frac{1}{2}\mu(\boldsymbol{\kappa}^T \boldsymbol{\kappa} - h^2), \tag{17}$$

where μ is an undetermined multiplier. The stationary point of the Lagrangian is obtained by solving the level-shifted Newton equation

$$\left(\boldsymbol{\xi}_m^{[2]} - \mu \mathbf{1}\right)\boldsymbol{\kappa}(\mu) = -\boldsymbol{\xi}_m^{[1]}. \tag{18}$$

In Ref. 37, it is shown that many stationary points may be found on the boundary, and that the one representing the minimum is uniquely determined, if μ is restricted to the interval $\mu < \epsilon_1$, where ϵ_1 is the lowest eigenvalue of the Hessian $\boldsymbol{\xi}_m^{[2]}$. We therefore need to determine the value μ for which the solution to Eq. (18) satisfies

$$\boldsymbol{\kappa}(\mu)^T \boldsymbol{\kappa}(\mu) = h^2; \qquad \mu < \epsilon_1. \tag{19}$$

Downloaded 15 Jun 2011 to 130.225.22.254. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/about/rights_and_permissions

The Newton or level-shifted Newton equations represent a set of linear equations which may be solved using an iterative algorithm. This requires that the gradient $\boldsymbol{\xi}_m^{[1]}$ and the linear transformation of the Hessian matrix on a trial vector $\boldsymbol{\xi}_m^{[2]}\boldsymbol{\kappa}$ can be determined. In Sec. II C we describe an efficient evaluation of these quantities.

The orbitals of the next iteration are obtained as

$$(|\tilde{p}\rangle|\tilde{q}\rangle...|\tilde{s}\rangle) = (|p\rangle|q\rangle...|s\rangle)\exp(-\boldsymbol{\kappa}), \tag{20}$$

where $\boldsymbol{\kappa}$ is determined either from the Newton or level-shifted Newton equation depending on the step that is taken.

Let us now describe how the trust-radius is updated. In iteration n, we have determined a step $\boldsymbol{\kappa}_n$ that generates the function value $(\xi_m)_{n+1}$ of iteration $n+1$. A prediction of this value may be obtained from the second-order expansion,

$$Q_m(\boldsymbol{\kappa}_n) = \left(\xi_m^{[0]}\right)_n + \boldsymbol{\kappa}_n^T\left(\boldsymbol{\xi}_m^{[1]}\right)_n + \frac{1}{2}\boldsymbol{\kappa}_n^T\left(\boldsymbol{\xi}_m^{[2]}\right)_n\boldsymbol{\kappa}_n, \tag{21}$$

where $(\xi_m)_{n+1}$ differs from $Q_m(\boldsymbol{\kappa}_n)$ through terms that are third and higher orders in $\boldsymbol{\kappa}_n$,

$$(\xi_m)_{n+1} = Q_m(\boldsymbol{\kappa}_n) + \mathcal{O}\left(\boldsymbol{\kappa}_n^3\right). \tag{22}$$

Let us consider the ratio between the observed and predicted change in the function ξ_m in iteration $n+1$,

$$r = \frac{(\xi_m)_{n+1} - \left(\xi_m^{[0]}\right)_n}{Q_m(\boldsymbol{\kappa}_n) - \left(\xi_m^{[0]}\right)_n} = 1 + \frac{\mathcal{O}\left(\boldsymbol{\kappa}_n^3\right)}{Q_m(\boldsymbol{\kappa}_n) - \left(\xi_m^{[0]}\right)_n}. \tag{23}$$

The closer r is to unity, the more quadratic the surface. We update the trust radius according to how quadratic the surface is

if $(r > 0.9)$,	**then**	h_{n+1}=1.2h_n;
if $(0.9>r>0.5)$,	**then**	h_{n+1}=h_n;
if $(0.5>r>0.2)$,	**then**	h_{n+1}=0.7h_n;
if $(0.2>r)$,	**then**	reject step, decrease trust-radius, and generate new step.

The above update of the trust-radius is conservative compared to usual update of the trust-region. This may increase the number of steps in the iterative procedure, but this is not important as there is little overhead involved in starting up a new iteration. For optimizing ξ_m it is most important to move safely on the function surface, in particular, because the lowest eigenvalue of the Hessian may change dramatically during the iterative procedure.

C. Determination of $\boldsymbol{\xi}_m^{[1]}$ and $\boldsymbol{\xi}_m^{[2]}\boldsymbol{\kappa}$

We will now describe how the gradient $\boldsymbol{\xi}_m^{[1]}$ and the linear transformation of the Hessian on a trial vector $\boldsymbol{\xi}_m^{[2]}\boldsymbol{\kappa}$ can be efficiently determined. We will identify $\boldsymbol{\xi}_m^{[1]}$ and $\boldsymbol{\xi}_m^{[2]}\boldsymbol{\kappa}$ from a linear expansion of $\partial\tilde{\xi}_m/\partial\kappa_{kl}$. First we write $\partial\tilde{\xi}_m/\partial\kappa_{kl}$ as

$$\frac{\partial\tilde{\xi}_m}{\partial\kappa_{kl}} = m\sum_p \tilde{\Omega}_p^{m-1}\frac{\partial\tilde{\Omega}_p}{\partial\kappa_{kl}}, \tag{24}$$

where $\tilde{\Omega}_p^m$ of Eq. (13) is defined only for positive integers. Since, we are interested in the linear expansion in $\boldsymbol{\kappa}$ in Eq. (24), we start out by determining the linear expansion of the individual terms in Eq. (24) using $\boldsymbol{\kappa} = \mathbf{0}$ as the expansion point,

$$\tilde{\Omega}_p^{m-1} = f_p^{[0]} + \boldsymbol{\kappa}^T\boldsymbol{f}_p^{[1]} + \mathcal{O}(\boldsymbol{\kappa}^2), \tag{25}$$

$$\frac{\partial\tilde{\Omega}_p}{\partial\kappa_{kl}} = \left(g_p^{[0]}\right)_{kl} + \left(\boldsymbol{\kappa}^T\boldsymbol{g}_p^{[1]}\right)_{kl} + \mathcal{O}(\boldsymbol{\kappa}^2), \tag{26}$$

$$f_p^{[0]} = \Omega_p^{m-1}, \tag{27}$$

$$\begin{aligned}\boldsymbol{\kappa}^T\mathbf{f}_p^{[1]} &= (m-1)\Omega_p^{m-2}\sum_x(\langle p|[\hat{\kappa},\hat{x}^2]|p\rangle \\ &\quad - 2\langle p|[\hat{\kappa},\hat{x}]|p\rangle\langle p|\hat{x}|p\rangle) \\ &= (m-1)\Omega_p^{m-2}\sum_x(2(\boldsymbol{\kappa}\mathbf{x}^2)_{pp} - 4(\boldsymbol{\kappa}\mathbf{x})_{pp}x_{pp}),\end{aligned} \tag{28}$$

$$\begin{aligned}\left(g_p^{[0]}\right)_{kl} &= \sum_x(2\langle p|[E_{kl},\hat{x}]^2|p\rangle - 4\langle p|[E_{kl},\hat{x}]|p\rangle\langle p|\hat{x}|p\rangle) \\ &= \sum_x\left(2\left(x_{lk}^2\delta_{kp} - x_{kl}^2\delta_{lp}\right) - 4\left(x_{lk}\delta_{kp} - x_{kl}\delta_{lp}\right)x_{pp}\right),\end{aligned} \tag{29}$$

$$\begin{aligned}(\boldsymbol{\kappa}^T\mathbf{g}_p^{[1]})_{kl} &= \sum_x(2\langle p|[E_{kl},[\hat{\kappa},\hat{x}^2]]|p\rangle \\ &\quad + 2\langle p|[\hat{\kappa},[E_{kl},\hat{x}^2]]|p\rangle \\ &\quad + 4\langle p|[E_{kl},\hat{x}]|p\rangle\langle p|[\hat{\kappa},\hat{x}]|p\rangle \\ &\quad - 4\langle p|[E_{kl},[\hat{\kappa},\hat{x}]]|p\rangle\langle p|\hat{x}|p\rangle \\ &\quad - 4\langle p|[[\hat{\kappa},E_{kl}],\hat{x}]]|p\rangle\langle p|\hat{x}|p\rangle) \\ &= 2\sum_x((\boldsymbol{\kappa}\mathbf{x}^2 - \mathbf{x}^2\boldsymbol{\kappa})_{lk}\delta_{kp} - (\boldsymbol{\kappa}\mathbf{x}^2 - \mathbf{x}^2\boldsymbol{\kappa})_{kl}\delta_{lp}) \\ &\quad - 8\sum_x(x_{lk}\delta_{kp} - x_{kl}\delta_{lp})(\boldsymbol{\kappa}\mathbf{x})_{pp} \\ &\quad - 4\sum_x((\boldsymbol{\kappa}\mathbf{x} - \mathbf{x}\boldsymbol{\kappa})_{lk}\delta_{kp} - (\boldsymbol{\kappa}\mathbf{x} - \mathbf{x}\boldsymbol{\kappa})_{kl}\delta_{lp})x_{pp} \\ &\quad + 1/2\sum_j\kappa_{jk}\left(g_p^{[0]}\right)_{jl} - 1/2\sum_j\kappa_{lj}\left(g_p^{[0]}\right)_{kj}.\end{aligned} \tag{30}$$

To obtain the last equality in Eq. (30) we have used the Jacobi identity for commutators.

The gradient $\boldsymbol{\xi}_m^{[1]}$ may now be identified as the term in Eq. (24) which is independent of $\boldsymbol{\kappa}$,

$$\begin{aligned}\left(\xi_m^{[1]}\right)_{kl} &= m\sum_p f_p^{[0]}\left(g_p^{[0]}\right)_{kl} \\ &= -2mQ_{kl}\Omega_l^{m-1} + 2m\Omega_k^{m-1}Q_{kl} \\ &\quad -4m\sum_x(\Omega_k^{m-1}x_{kk}x_{kl} - x_{kl}x_{ll}\Omega_l^{m-1}),\end{aligned} \tag{31}$$

Downloaded 15 Jun 2011 to 130.225.22.254. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/about/rights_and_permissions

where we have introduced

$$\mathbf{Q} = \mathbf{x}^2 + \mathbf{y}^2 + \mathbf{z}^2. \tag{32}$$

Introducing the odd component of a matrix $\mathbf{A}$ as

$$\{\mathbf{A}\}^{\mathrm{o}} = \mathbf{A} - \mathbf{A}^{\dagger}, \tag{33}$$

we may write the gradient in matrix notation as

$$\boldsymbol{\xi}_m^{[1]} = 2m\{\mathrm{dia}(\boldsymbol{\Omega}^{m-1})\mathbf{Q}\}^{\mathrm{o}} - 4m\left\{\sum_x \mathrm{dia}(\mathbf{x})\mathrm{dia}(\boldsymbol{\Omega}^{m-1})\mathbf{x}\right\}^{\mathrm{o}}, \tag{34}$$

where $\mathrm{dia}(\mathbf{x})$ means that the diagonal elements are taken out of the matrix $\mathbf{x}$. Note that for $m = 1$ the first term in Eq. (34) vanishes, because $\mathbf{Q}$ is a symmetric matrix. This is as expected since the optimization of the Boys function in Eq. (1) does not depend on $\mathbf{Q}$.

The linear transformation of the Hessian on a trial vector $\boldsymbol{\kappa}$ may similarly be identified from the terms in Eq. (24) which are linear in $\boldsymbol{\kappa}$,

$$\begin{aligned}
\left(\boldsymbol{\xi}_m^{[2]}\boldsymbol{\kappa}\right)_{kl} &= \sum_p \left(mf_p^{[0]}\left(\boldsymbol{\kappa}^T\mathbf{g}_p^{[1]}\right)_{kl} + m\boldsymbol{\kappa}^T\mathbf{f}_p^{[1]}\left(g_p^{[0]}\right)_{kl}\right) \\
&= 2m\Omega_k^{m-1}(\boldsymbol{\kappa}\mathbf{Q} - \mathbf{Q}\boldsymbol{\kappa})_{lk} - 2m\Omega_l^{m-1}(\boldsymbol{\kappa}\mathbf{Q} - \mathbf{Q}\boldsymbol{\kappa})_{kl} \\
&\quad - 8m\Omega_k^{m-1}\sum_x x_{lk}(\boldsymbol{\kappa}\mathbf{x})_{kk} + 8m\Omega_l^{m-1}\sum_x x_{kl}(\boldsymbol{\kappa}\mathbf{x})_{ll} \\
&\quad - 4m\Omega_k^{m-1}\sum_x (\boldsymbol{\kappa}\mathbf{x} - \mathbf{x}\boldsymbol{\kappa})_{lk}x_{kk} \\
&\quad + 4m\Omega_l^{m-1}\sum_x (\boldsymbol{\kappa}\mathbf{x} - \mathbf{x}\boldsymbol{\kappa})_{kl}x_{ll} \\
&\quad + 1/2\sum_j \left(\kappa_{jk}\left(\xi_m^{[1]}\right)_{jl} - \kappa_{lj}\left(\xi_m^{[1]}\right)_{kj}\right) \\
&\quad + 4m(m-1)\Omega_k^{m-2}(\boldsymbol{\kappa}\mathbf{Q})_{kk}Q_{lk} \\
&\quad - 4m(m-1)\Omega_l^{m-2}(\boldsymbol{\kappa}\mathbf{Q})_{ll}Q_{kl} \\
&\quad - 8m(m-1)\Omega_k^{m-2}\sum_x (\boldsymbol{\kappa}\mathbf{x})_{kk}x_{kk}Q_{lk} \\
&\quad + 8m(m-1)\Omega_l^{m-2}\sum_x (\boldsymbol{\kappa}\mathbf{x})_{ll}x_{ll}Q_{kl} \\
&\quad - 8m(m-1)\Omega_k^{m-2}(\boldsymbol{\kappa}\mathbf{Q})_{kk}\sum_y y_{lk}y_{kk} \\
&\quad + 8m(m-1)\Omega_l^{m-2}(\boldsymbol{\kappa}\mathbf{Q})_{ll}\sum_y y_{kl}y_{ll} \\
&\quad + 16m(m-1)\Omega_k^{m-2}\sum_x (\boldsymbol{\kappa}\mathbf{x})_{kk}x_{kk}\sum_y y_{lk}y_{kk} \\
&\quad - 16m(m-1)\Omega_l^{m-2}\sum_x (\boldsymbol{\kappa}\mathbf{x})_{ll}x_{ll}\sum_y y_{kl}y_{ll}.
\end{aligned} \tag{35}$$

Using the matrix notation of Eq. (33) we may write the linear transformation of the Hessian on a trial vector as

$$\begin{aligned}
\boldsymbol{\xi}_m^{[2]}\boldsymbol{\kappa} &= 2m\{\mathrm{dia}(\boldsymbol{\Omega}^{m-1})(\boldsymbol{\kappa}\mathbf{Q} - \mathbf{Q}\kappa)\}^{\mathrm{o}} \\
&\quad - 8m\{\mathrm{dia}(\boldsymbol{\Omega}^{m-1})\sum_x \mathrm{dia}(\boldsymbol{\kappa}\mathbf{x})\mathbf{x}\}^{\mathrm{o}} \\
&\quad - 4m\{\mathrm{dia}(\boldsymbol{\Omega}^{m-1})\sum_x \mathrm{dia}(\mathbf{x})(\boldsymbol{\kappa}\mathbf{x} - \mathbf{x}\boldsymbol{\kappa})\}^{\mathrm{o}} \\
&\quad - 1/2\left\{\boldsymbol{\kappa}\boldsymbol{\xi}_m^{[1]}\right\}^{\mathrm{o}} \\
&\quad + 4m(m-1)\{\mathrm{dia}(\boldsymbol{\Omega}^{m-2})\mathrm{dia}(\boldsymbol{\kappa}\mathbf{Q})\mathbf{Q}\}^{\mathrm{o}} \\
&\quad - 8m(m-1)\{\mathrm{dia}(\boldsymbol{\Omega}^{m-2})\sum_x \mathrm{dia}(\boldsymbol{\kappa}\mathbf{x})\mathrm{dia}(\mathbf{x})\mathbf{Q}\}^{\mathrm{o}} \\
&\quad - 8m(m-1)\{\mathrm{dia}(\boldsymbol{\Omega}^{m-2})\mathrm{dia}(\boldsymbol{\kappa}\mathbf{Q})\sum_y \mathrm{dia}(\mathbf{y})\mathbf{y}\}^{\mathrm{o}} \\
&\quad + 16m(m-1)\left\{\mathrm{dia}(\boldsymbol{\Omega}^{m-2})\right. \\
&\quad \left.\times\left(\sum_x \mathrm{dia}(\boldsymbol{\kappa}\mathbf{x})\mathrm{dia}(\mathbf{x})\right)\left(\sum_y \mathrm{dia}\left(\mathbf{y}\right)\mathbf{y}\right)\right\}^{\mathrm{o}}.
\end{aligned} \tag{36}$$

Note that the linear transformation in Eq. (36) contains five matrix multiplications $\boldsymbol{\kappa}\mathbf{Q}$, $\boldsymbol{\kappa}\mathbf{x}$, $\boldsymbol{\kappa}\mathbf{y}$, $\boldsymbol{\kappa}\mathbf{z}$, and $\boldsymbol{\kappa}\boldsymbol{\xi}_m^{[1]}$.

When solving the Newton or level-shifted Newton equation, the linear equations may be preconditioned using the diagonal of the Hessian matrix,

$$\begin{aligned}
\left(\xi_m^{[2]}\right)_{kl,kl} &= 2m(Q_{kk} - Q_{ll})\left(\Omega_l^{m-1} - \Omega_k^{m-1}\right) \\
&\quad - 8m\sum_x (x_{lk})^2\left(\Omega_k^{m-1} + \Omega_l^{m-1}\right) \\
&\quad + 4m\Omega_k^{m-1}\sum_x ((x_{kk})^2 - x_{ll}x_{kk}) \\
&\quad + 4m\Omega_l^{m-1}\sum_x ((x_{ll})^2 - x_{ll}x_{kk}) \\
&\quad + 4m(m-1)(Q_{lk})^2\left(\Omega_k^{m-2} + \Omega_l^{m-2}\right) \\
&\quad - 16m(m-1)Q_{lk}\sum_x x_{lk}\left(\Omega_k^{m-2}x_{kk} + \Omega_l^{m-2}x_{ll}\right) \\
&\quad + 16m(m-1)\Omega_k^{m-2}\left(\sum_x x_{lk}x_{kk}\right)\left(\sum_y y_{lk}y_{kk}\right) \\
&\quad + 16m(m-1)\Omega_l^{m-2}\left(\sum_x x_{lk}x_{ll}\right)\left(\sum_y y_{kl}y_{ll}\right).
\end{aligned} \tag{37}$$

We note that the gradient term in the linear transformation in Eq. (36) does not give a contribution to the diagonal Hessian elements because $\boldsymbol{\xi}_m^{[1]}$ is an antisymmetric matrix.

III. RESULTS

In this section, we give examples of local occupied and virtual HF orbitals obtained using the localization function $\boldsymbol{\xi}_m$ in Eq. (8). We present localized orbitals for the following molecules: Superbenzene ($C_{24}H_{12}$) with a B3LYP/cc-pVTZ optimized geometry; Buckminster-fullerene (C_{60}) with

Downloaded 15 Jun 2011 to 130.225.22.254. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/about/rights_and_permissions

a B3LYP/6-31G(d,p) optimized geometry; and the $I27_{SS}$ domain of the titin protein (392 atoms) with a BP86/6-31G optimized geometry (but with 6-31G* basis functions on the three sulfur atoms and on two carbon atoms as described in Ref. 38). We consider calculations using Dunning's correlation-consistent basis sets,[39] cc-pVXZ for X=D,T,Q. For C_{60} we only report calculations using the cc-pVDZ and cc-pVTZ basis sets and for the titin fragment only for the cc-pVDZ basis set. We investigate the effect of increasing m in Eq. (8) and of increasing the cardinal number. The difficulties encountered when optimizing the locality function $\boldsymbol{\xi}_m$ depend on whether a set of occupied or a set of unoccupied HF orbitals are considered. We discuss the origin of this difference. In any case, the computational cost of the localization procedure is negligible compared to the total HF calculation.

A. Maximum orbital spread

The maximum orbital spread (MOS) for a set of orbitals $\mathcal{S}$ is given by

$$\text{MOS} = \max_{p\in\mathcal{S}} \sigma_p, \tag{38}$$

where σ_p is given in Eq. (6). We use the MOS as a measure of the locality of the set $\mathcal{S}$, which refers either to the set of occupied orbitals or to the set of virtual orbitals.

The MOSs are presented for superbenzene (Table I), C_{60} (Table II), and the titin fragment (Table III) for the occupied (Occ) and virtual (Virt) HF orbitals when $\boldsymbol{\xi}_m$ is optimized for $m = 1, 2, \ldots, 10$. For the titin fragment only $m = 1, 2, \ldots, 5$ results are reported. In Fig. 1 we present a graphical illustration of the occupied and virtual HF orbitals with maximum orbitals spreads for superbenzene, C_{60}, and the titin fragment for $m = 2$.

1. Different molecules and different cardinal numbers

For the occupied HF orbitals, the MOS for $m = 1$ is similar for X = D and X = T, while for X = Q there is a significant increase. For example, for superbenzene with $m = 1$ the MOS is 2.29 (cc-pVDZ), 2.29 (cc-pVTZ), and 3.01 (cc-pVQZ). When m increases the MOS of the occupied orbitals decreases and becomes similar to all cardinal numbers. For superbenzene for $m = 10$ we thus obtain 2.13 (cc-pVDZ), 2.13 (cc-pVTZ), and 2.15 (cc-pVQZ) for the MOS for the occupied orbitals. The similarity between the MOSs for the occupied orbitals reflects that the occupied orbitals are described well at the cc-pVDZ level and that only small changes are introduced in the occupied orbitals when the cardinal number is increased. For larger m values the MOSs become similar, even for molecules with rather different electronic structures. For example, for $m = 5$ the MOS for the occupied orbitals in the cc-pVDZ calculations is 2.17, 2.10, and 2.13 for superbenzene, C_{60}, and the titin fragment, respectively.

For the unoccupied HF orbitals, the MOS varies somewhat with the cardinal number. For $m = 1$ the MOS increases with the cardinal number and outliers may be encountered. For example, for superbenzene the MOS for the unoccupied HF orbitals is 3.00 (cc-pVDZ), 3.26 (cc-pVTZ), and 7.61 (cc-pVQZ), where the latter clearly is an outlier. For larger m

TABLE I. Maximum orbital spreads (a.u.) for localized superbenzene orbitals using different penalty exponents m. Also shown are maximum orbital spreads (a.u.) for LCM orbitals, PAO, and canonical molecular orbitals (CMO).

	cc-pVDZ		cc-pVTZ		cc-pVQZ	
	Occ	Virt	Occ	Virt	Occ	Virt
1	2.288	3.004	2.292	3.261	3.012	7.614
2	2.253	2.717	2.256	2.447	2.258	2.560
3	2.226	2.263	2.228	2.317	2.239	2.337
4	2.192	2.245	2.193	2.113	2.208	2.241
5	2.170	2.234	2.171	2.064	2.192	2.177
6	2.156	2.224	2.157	2.050	2.180	2.122
7	2.146	2.213	2.146	2.040	2.171	2.079
8	2.138	2.202	2.138	2.032	2.163	2.041
9	2.131	2.194	2.132	2.025	2.157	2.021
10	2.126	2.189	2.127	2.020	2.151	2.007
LCM	3.300	4.152	3.300	4.553	3.444	5.747
PAO	...	3.550	...	3.551	...	3.844
CMO	7.458	10.737	7.452	11.551	7.452	12.160

TABLE II. Maximum orbital spreads (a.u.) for localized C_{60} orbitals using different penalty exponents m. Also shown are maximum orbital spreads (a.u.) for PAO and CMO.

	cc-pVDZ		cc-PVTZ	
	Occ	Virt	Occ	Virt
1	2.281	2.868	2.285	3.212
2	2.156	2.426	2.159	2.469
3	2.128	2.349	2.131	2.310
4	2.114	2.311	2.118	2.235
5	2.100	2.288	2.101	2.189
6	2.088	2.274	2.089	2.160
7	2.080	2.263	2.081	2.140
8	2.073	2.255	2.074	2.124
9	2.068	2.249	2.069	2.112
10	2.064	2.244	2.065	2.103
PAO	...	3.469	...	3.497
CMO	7.211	9.257	7.222	10.225

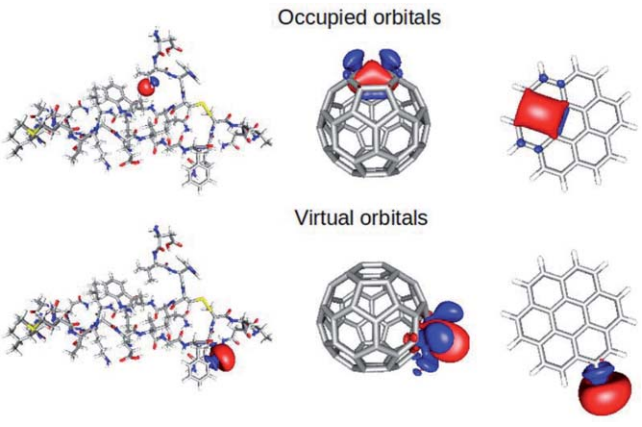

FIG. 1. Least local orbitals for the occupied and virtual orbital spaces for the $I27_{SS}$ domain of the titin protein (left), C_{60} (middle), and superbenzene (right) plotted with orbital amplitude values of 0.03 (a.u.). The calculations were carried out using the cc-pVDZ basis and a power exponent $m = 2$.

Downloaded 15 Jun 2011 to 130.225.22.254. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/about/rights_and_permissions

TABLE III. Maximum orbital spreads (a.u.) for localized orbitals for the $I27_{SS}$ domain of the titin protein using different penalty exponents m. Also shown are maximum orbital spreads (a.u.) for PAO and CMO.

	cc-pVDZ	
	Occ	Virt
1	2.245	3.483
2	2.188	2.644
3	2.163	2.519
4	2.153	2.467
5	2.133	2.415
PAO	...	4.710
CMO	24.754	26.334

values the MOS decreases and the outliers disappear. It is also seen that for larger m values the MOS may become smaller for the virtual orbitals when the cardinal number is increased. For example, for C_{60} and $m = 10$ the MOS is 2.24 (cc-pVDZ) and 2.10 (cc-pVTZ). To understand this, recall that when the cardinal number is increased from X to $X+1$, we add atomic basis functions which have a significantly smaller orbital spread, than the localized molecular HF orbitals for cardinal number X. The added atomic basis functions therefore have the flexibility to make the MOS smaller for cardinal number $X+1$ than for X.

It is also worth noticing that for larger m values the difference between the MOS for the occupied and for the virtual orbitals decreases. In fact, for larger cardinal numbers the virtual orbitals may have a smaller MOS than the occupied orbitals. For example, for the cc-pVQZ superbenzene calculation with $m = 10$, the MOS is 2.15 for the occupied orbitals and 2.01 for the virtual orbitals.

As it was also the case for the occupied orbitals, the locality of the virtual HF orbitals is nearly system independent for larger m values and for a given basis set. For example, for the cc-pVDZ basis and $m = 5$ the MOS for the virtual orbitals is 2.23, 2.29, and 2.42 for superbenzene, C_{60}, and the titin fragment, respectively.

2. Comparing local virtual orbitals with PAOs

In Tables I, II, and III, we have also reported the MOSs for the PAOs spanning the unoccupied HF orbital space and (only Table I) for the occupied and unoccupied LCM orbitals. The (redundant) PAOs in Tables I, II, and III were constructed by projecting out the occupied orbital space from the atomic basis functions expressed in the canonical atomic (CA) basis, which enters level 1 of the three-level optimization procedure.[35,36] For completeness we mention that the MOSs for the PAOs in the standard AO basis and in the CA basis are very similar (they differ by at most 0.1 a.u. for the cases we have investigated).

The MOS for the LCM orbitals is significantly larger than for our localized orbitals, while the PAO values are in between the values for the unoccupied LCM orbitals and our localized unoccupied orbitals. In general, the locality of the PAOs seems to decrease when the size of the molecular system is increased. For the cc-pVDZ basis, the MOS for the PAOs is 3.55 for superbenzene, 3.47 for C_{60}, and 4.71 for the titin fragment. For the titin fragment in the cc-pVDZ basis, the MOS for the least local PAO (4.71) is almost twice the MOS for our least local virtual orbital in the $m = 2$ case (2.64). The fact that the MOS is a meaningful measure of locality is supported by Fig. 2, where these two orbitals are displayed for amplitude values of 0.03, 0.003, and 0.0003. The displayed orbitals clearly demonstrate that the least local PAO is significantly less local than our least local virtual orbital for $m = 2$.

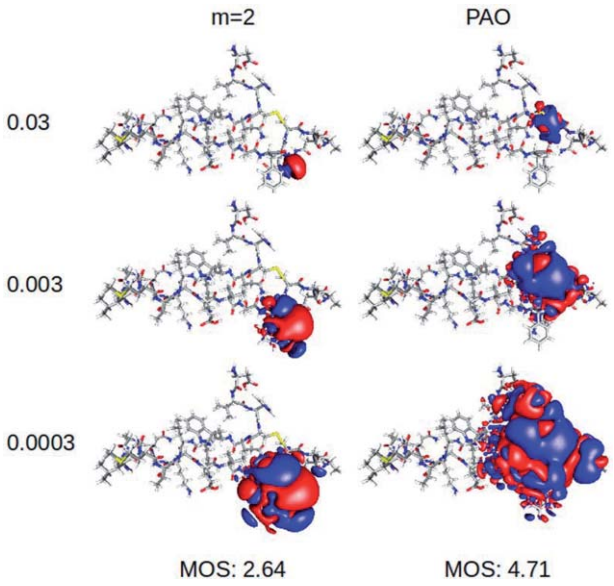

FIG. 2. Least local virtual orbital for power exponent $m = 2$ (left), and least local PAO (right) plotted for different orbital amplitude values 0.03, 0.003, and 0.0003 (a.u.). The calculations were carried out on the $I27_{SS}$ domain of the titin protein using the cc-pVDZ basis.

B. Detailed analysis of the local orbitals for superbenzene

To understand what is happening when the penalty for delocalized orbitals is increased we have in Fig. 3 displayed the orbital spreads for the individual occupied orbitals in a cc-pVQZ calculation for superbenzene for $m = 1, 2, 5$, and 10. The orbitals are ordered starting with 24 core orbitals, followed by 54 valence orbitals. The orbital spreads for the core orbitals are the same for all m values because the core orbitals have been taken out of the localization function. For the valence orbitals a wide distribution of orbital spreads is observed for $m = 1$ – the minimum and maximum orbital spreads are 1.41 and 3.01, respectively. When m is increased, the orbital spread distribution gets much more narrow, and for $m = 10$ the minimum and maximum orbital spreads are 1.68 and 2.15, respectively. Thus, when increasing m, the valence orbitals with the largest orbital spreads become more local, at the expense that the orbitals with the smallest orbital spreads are delocalized slightly. This conclusion is supported by the graphical representations of the orbitals with minimum and maximum orbital spreads in Fig. 3. Clearly, when m is increased, the orbital with the largest orbital spread becomes more compact, while the orbital with the smallest orbital spread expands.

Downloaded 15 Jun 2011 to 130.225.22.254. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/about/rights_and_permissions

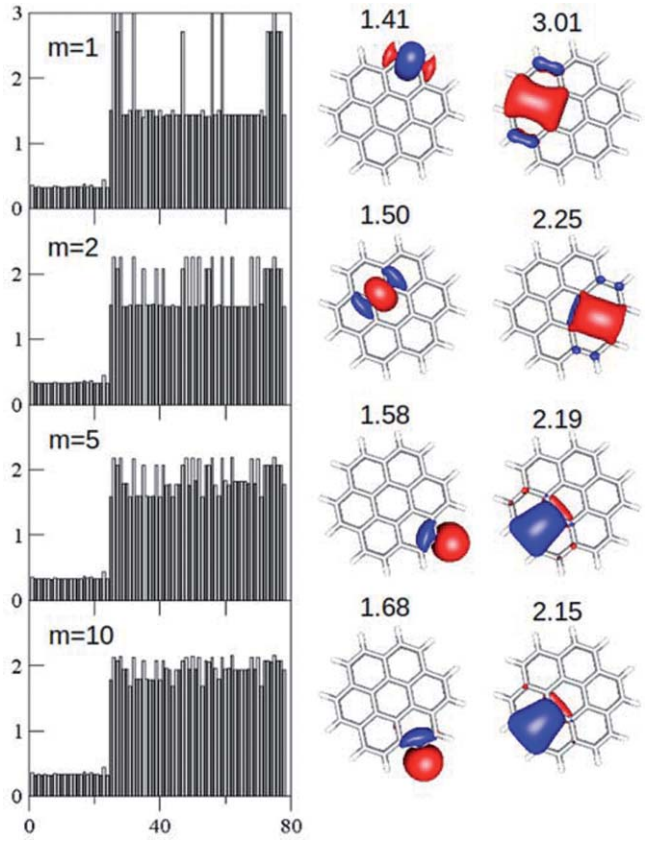

FIG. 3. Occupied orbitals obtained with power exponents $m = 1, 2, 5$, and 10 in a cc-pVQZ calculation on the superbenzene molecule. The orbitals were plotted using orbital amplitude values of 0.03 (a.u.). Left column: Orbital spreads (a.u.) against orbital number for all (core+valence) occupied orbitals. Middle column: *Most local* occupied valence orbital and associated orbital spread. Right column: *Least local* occupied valence orbital and associated orbital spread.

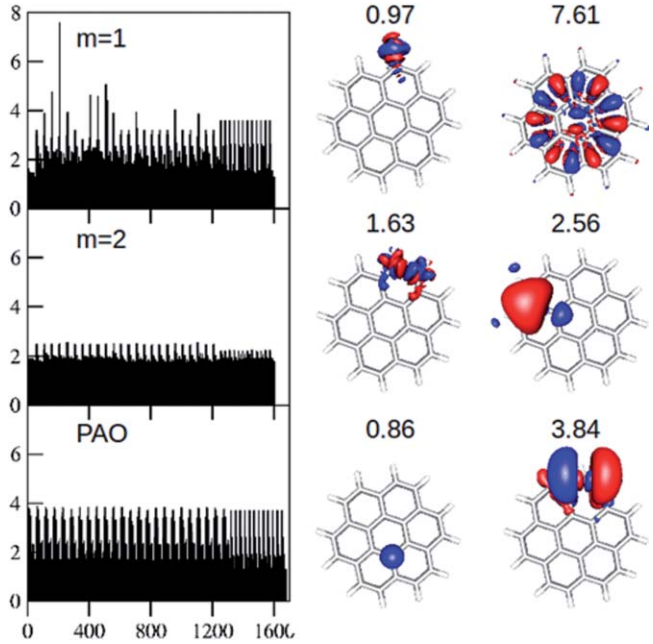

FIG. 4. Virtual orbitals obtained with power exponents $m = 1$ and $m = 2$ and the PAOs in a cc-pVQZ calculation on the superbenzene molecule. The orbitals were plotted using orbital amplitude values of 0.03 (a.u.). Left column: Orbital spreads (a.u.) against orbital number for all virtual orbitals. Middle column: *Most local* virtual orbital or PAO and associated orbital spread. Right column: *Least local* virtual orbital or PAO and associated orbital spread.

The virtual HF orbitals follow the same pattern as the occupied valence orbitals. In Fig. 4 we have displayed similar plots for the virtual orbitals, restricting ourselves here to $m = 1$ and $m = 2$. It is seen that outliers occur for $m = 1$. In particular, Fig. 4 reveals the presence of one extreme outlier orbital with an orbital spread of 7.61, which is basically spread out over the full molecular system. By increasing the penalty to $m = 2$ the outliers disappear, and the orbital spread distribution is narrowed, such that all virtual orbitals get an orbital spread of about 2. As it was also discussed for the occupied orbitals above, Fig. 4 illustrates that the removal of virtual outlier orbitals when m is increased happens at the expense that the most local virtual orbitals get delocalized. For example, when the maximum orbital spread is decreased from 7.61 to 2.56 by increasing m, the minimum orbital spread is increased from 0.97 to 1.63.

In Fig. 4 we have also displayed the orbital spreads for all PAOs and plotted the PAOs with the smallest (0.86) and largest (3.84) orbital spreads. The orbital spreads for the PAOs are in general significantly larger than the orbital spreads for the virtual orbitals with $m = 2$. It is also seen that the distribution of orbital spreads is much more homogeneous for the virtual orbitals with $m = 2$ than for the PAOs.

C. Details about the optimization of the locality functions

Optimization of localization functions has previously been performed using a Jacobi sweep of iterations. However, this method is too simplistic for minimizing functions of the complexity of Eq. (8). We have instead used the trust-region method of Fletcher[37] to minimize $\boldsymbol{\xi}_m$. In the trust-region method, both gradient and Hessian information is used, and convergence to a minimum is guaranteed. Using the trust-region method, Eq. (8) can be minimized for an arbitrary m, for both the occupied and the unoccupied HF orbitals. The minimization is most difficult for the unoccupied set of HF orbitals, where many very large negative Hessian eigenvalues are encountered in the initial iterations. During the minimization, these negative eigenvalues vanish, and quadratic convergence is obtained in the last iterations. For the occupied HF orbitals only small negative Hessian eigenvalues are encountered in the initial iterations and these disappear fast during the optimization. This is probably the reason that a Jacobi sweep of iterations has been able to optimize the Boys function for the occupied orbitals.

As starting orbitals we have used a simplified version of the LCM orbitals. For these simplified LCM orbitals the expensive procedure, where all possible orbital orderings of the orthonormal orbitals are investigated, has been replaced by a cheap and simple reordering scheme based on the overlap between the orthonormal AOs and the set of occupied orbitals. The simplified LCM orbitals are not nearly as local as when the full orbital ordering has been carried out (as it was done for the LCM orbitals in Table I). In general, however, the

Downloaded 15 Jun 2011 to 130.225.22.254. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/about/rights_and_permissions

simplified LCM orbitals are significantly more local than the canonical orbitals, and therefore they provide a better starting guess for generating local orbitals.

IV. SUMMARY AND PERSPECTIVE

We have demonstrated that a set of local orthonormal HF orbitals can be obtained both for the occupied and virtual orbital spaces. The local HF orbitals are obtained by setting up localization functions where not only the set of occupied (or virtual) HF orbitals on average are most local, but where a penalty is also imposed on delocalized orbitals. As a measure of locality we have used the orbital variance, and penalties are introduced by using powers of the orbital variance as a locality measure. Our general localization function becomes the Boys function for $m = 1$.

The Boys localization function has previously been successfully optimized using a Jacobi sweep of iterations for the occupied HF orbitals. However, the Jacobi sweep has failed to give local virtual HF orbitals. We have demonstrated that using the trust-region method the optimization of the Boys function and its generalizations can be performed both for the occupied and virtual orbital spaces.

Numerical examples have shown that the optimization of the Boys function may lead to outlier orbitals with orbitals spreads that are much larger than for the rest of the orbitals. These outliers disappear when penalties are imposed on large orbital spreads, and for larger penalties the maximum orbital spread becomes similar to the occupied and virtual HF orbitals. In fact, increasing the cardinal number for Dunning's correlation consistent basis sets for larger penalties leads to virtual HF orbitals, which are more local than the occupied HF orbitals. The numerical results also demonstrate that for larger penalties the locality of the MOs is very similar to molecules with rather different electronic structures. Even the MOs of superbenzene – which one would intuitively expect to have a highly delocalized electronic structure – can be made highly local.

Our local virtual HF orbitals are significantly more local than the PAOs that often have been used in local correlation wave function calculations to span the virtual orbital space. Furthermore, our local virtual orbitals are non-redundant and orthonormal, in contrast to the PAOs. It is important for local correlation methods to have highly local HF orbitals for both the occupied and virtual orbital spaces, and that *no* outlier orbitals are encountered, since outliers impede an efficient description of local electron correlation effects. For example, a virtual outlier orbital would have a non-vanishing charge distribution with very many local occupied orbitals – e.g., the virtual outlier with orbital spread 7.61 in Fig. 4 has a non-vanishing charge distribution with all occupied orbitals. This would prohibit an efficient use of orbital locality in local correlated methods. Our local HF orbitals (for $m > 1$) thus appear to be a good candidate for local correlation methods.

ACKNOWLEDGMENTS

This work has been supported by the Lundbeck Foundation.

[1] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).
[2] S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).
[3] J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960).
[4] S. F. Boys, "Localized orbitals and localized adjustment functions," in *Quantum Theory of Atoms, Molecules and Solid State*, edited by P.-O. Löwdin (Academic, New York, 1966), p. 253.
[5] C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).
[6] C. Edmiston and K. Ruedenberg, J. Chem. Phys. **43**, S97 (1965).
[7] C. Edmiston and K. Ruedenberg, "Localized atomic and molecular orbitals. III," in *Quantum Theory of Atoms, Molecules and Solid State*, edited by P.-O. Löwdin (Academic, New York, 1966), p. 263.
[8] D. A. Kleier, T. A. Halgren, J. John, H. Hall, and W. N. Lipscomb, J. Chem. Phys. **61**, 3905 (1974).
[9] W. von Niessen, J. Chem. Phys. **56**, 4290 (1972).
[10] V. Magnasco and A. Perico, J. Chem. Phys. **47**, 971 (1967).
[11] J. Pipek, Int. J. Quantum Chem. **36**, 487 (1989).
[12] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).
[13] D. Maynau, S. Evangelisti, N. Guihéry, C. J. Calzado, and J.-P. Malrieu, J. Chem. Phys. **116**, 10060 (2002).
[14] C. Angeli, C. J. Calzado, R. Cimiraglia, S. Evangelisti, N. Guihéry, T. Leininger, J.-P. Malrieu, D. Maynau, J. V. P. Ruiz, and M. Sparta, Mol. Phys. **101**, 1389 (2003).
[15] F. Aquilante, T. B. Pedersen, A. S. de Merás, and H. Koch, J. Chem. Phys. **125**, 174101 (2006).
[16] J. E. Subotnik, A. D. Dutoi, and M. Head-Gordon, J. Chem. Phys. **123**, 114108 (2005).
[17] S. Saebø and P. Pulay, J. Chem. Phys. **86**, 914 (1987).
[18] S. Saebø and P. Pulay, J. Chem. Phys. **88**, 1884 (1988).
[19] S. Saebø and P. Pulay, Annu. Rev. Phys. Chem. **44**, 213 (1993).
[20] C. Hampel and H.-J. Werner, J. Chem. Phys. **104**, 6286 (1996).
[21] M. Schütz, G. Hetzer, and H.-J. Werner, J. Chem. Phys. **111**, 5691 (1999).
[22] G. Hetzer, M. Schütz, H. Stoll, and H.-J. Werner, J. Chem. Phys. **113**, 9443 (2000).
[23] M. Schütz, J. Chem. Phys. **113**, 9986 (2000).
[24] M. Schütz and H.-J. Werner, J. Chem. Phys. **114**, 661 (2001).
[25] P. E. Maslen, M. S. Lee, and M. Head-Gordon, Chem. Phys. Lett. **319**, 205 (2000).
[26] S. Li, J. Ma, and Y. Jiang, J. Comput. Chem. **23**, 237 (2001).
[27] W. Li, P. Piecuch, J. R. Gour, and S. Li, J. Chem. Phys. **131**, 114109 (2009).
[28] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 8644 (2010).
[29] H. Stoll, Chem. Phys. Lett. **191**, 548 (1992).
[30] J. Friedrich, M. Hanrath, and M. Dolg, J. Chem. Phys. **126**, 154110 (2007).
[31] O. Christiansen, P. Manninen, P. Jørgensen, and J. Olsen, J. Chem. Phys. **124**, 084103 (2006).
[32] V. Weijo, P. Manninen, P. Jørgensen, O. Christiansen, and J. Olsen, J. Chem. Phys. **127**, 074106 (2007).
[33] M. Ziółkowski, B. Jansík, P. Jørgensen, and J. Olsen, J. Chem. Phys. **131**, 124112 (2009).
[34] L. Cederbaum, J. Schirmer, and H.-D. Meyer, J. Phys. A **22**, 2427 (1989).
[35] B. Jansík, S. Høst, M. P. Johansson, P. Jørgensen, J. Olsen, and T. Helgaker, J. Chem. Theory Comput. **5**, 1027 (2009).
[36] B. Jansík, S. Høst, M. P. Johansson, J. Olsen, P. Jørgensen, and T. Helgaker, Phys. Chem. Chem. Phys. **11**, 5805 (2009).
[37] R. Fletcher, *Practical Methods of Optimization* (Wiley, Chichester, 1990).
[38] S. Reine, A. Krapp, M. F. Iozzi, V. Bakken, T. Helgaker, F. Pawłowski, and P. Sałek, J. Chem. Phys. **133**, 044102 (2010).
[39] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).

Downloaded 15 Jun 2011 to 130.225.22.254. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/about/rights_and_permissions