Orbital localization using fourth central moment minimization

Ida-Marie Høyvik, Branislav Jansik, and Poul Jørgensen

Orbital localization using fourth central moment minimization

Ida-Marie Høyvik,[1] Branislav Jansik,[2] and Poul Jørgensen[1]

[1]qLEAP Center for Theoretical Chemistry, Department of Chemistry, Aarhus University, Langelandsgade 140, Aarhus 8000, Denmark
[2]VSB-TUO, 17. listopadu 15/2172, 798 33 Ostrava-Poruba, Czech Republic

We present a new orbital localization function based on the sum of the fourth central moments of the orbitals. To improve the locality, we impose a power on the fourth central moment to act as a penalty on the least local orbitals. With power two, the occupied and virtual Hartree-Fock orbitals exhibit a more rapid tail decay than orbitals from other localization schemes, making them suitable for use in local correlation methods. We propose that the standard orbital spread (the square root of the second central moment) and fourth moment orbital spread (the fourth root of the fourth central moment) are used as complementary measures to characterize the locality of an orbital, irrespective of localization scheme. © 2012 American Institute of Physics. [http://dx.doi.org/10.1063/1.4769866]

I. INTRODUCTION

It is a basic assumption in local correlation methods that the wave function is expanded in a set of local orbitals. However, which criteria a set of orbitals have to fulfill in order to be characterized as local is not clear. In one of the popular textbooks in computational chemistry,[1] local orbitals are defined as orbitals "that are spatially confined to a relatively small volume, and therefore clearly displaying which atoms are bonded and furthermore have the property of being approximately constant between structurally similar units in different molecules." Locality is here intimately associated with molecular bonding and connected to the structural similarity of the chemical units in different molecules. In the Pipek–Mezey orbital localization scheme,[2,3] the idea is to reduce the number of atomic sites for a bonding orbital. The close connection between locality and chemical bonding is widely accepted despite the fact that molecular bonding is associated with the nodal structure of an orbital and with particular sets of orbitals, e.g., HOMO-LUMO gap and Koopman's theorem for canonical Hartree–Fock (HF) orbitals.

In the context of local correlation methods, local orbitals are introduced with the sole purpose of providing a local basis where the short range inter-electronic interactions (Coulomb holes and dispersion forces) can be described efficiently. For such a purpose, it is only important to use orbitals that are confined to a small volume in space, and their bonding properties are irrelevant. The problem is how to find appropriate criteria for determining and characterizing local orbitals.

A vast number of localization schemes (e.g., Refs. 3–13) have been proposed since Lennard-Jones and Pople[14] published an article on localized occupied orbitals and their physical meaning in 1950. The most used localization schemes are the ones of Boys,[4,5] Edmiston and Ruedenberg,[6] and Pipek and Mezey.[2,3] In solid state physics, the Boys and Pipek–Mezey schemes have been employed to obtain maximally localized Wannier orbitals.[15,16] The use of the Boys, Pipek–Mezey, and Edmiston–Ruedenberg schemes to obtain a local virtual orbital space has been very limited, due to the failure of the Jacobi sweeps of iterations[6] to optimize the functions for the virtual space. Projected atomic orbitals (PAOs), which constitute a redundant non-orthogonal set, have been used in some local correlation methods[17–19] in the absence of local virtual HF orbitals. Recently, Høyvik et al.[20] showed that a set of local, non-redundant orthonormal virtual orbitals can be obtained by using the trust region optimization of the localization functions.

The Boys scheme minimizes the sum of the second central moments (variances) of the orbitals. For simplicity, we will refer to the second central moment as the second moment. The second moment of an orbital characterizes the deviation of its density from its average position, and hence measures where the bulk of the orbital density is located. The Boys localization thus produces a set of orbitals, which on average are local, but the set may contain some orbitals, which have a large second moment—so-called outliers. For a set of orbitals, it is the locality of the least local orbital ("the weakest link") that defines the locality of the set. Jansik et al.[12] showed how outliers could be removed by imposing a power on the second moments, acting as a penalty for orbitals with large second moments.

When minimizing the second moment, no considerations are given to the shape of the orbital, except that the bulk of the orbital is concentrated in a small volume of space. However, a small second moment may be obtained because of an acute peak and a thick tail or because of a broad peak and a thin tail. In local correlation methods, local orbitals must have a thin tail since only then is the orbital really restricted to a small volume in space. In this paper, we describe how local orbitals with a fast tail decay may be obtained by minimizing the sum of the fourth moments for a set of orbitals. To avoid outliers in the fourth moment minimization, we further suggest to impose a power on the fourth moment in accordance with the power on the second moment as discussed in Ref. 12.

The second moment of an orbital is used to characterize the locality of an orbital since it reflects the spatial confinement of the bulk of the orbital. In practice, the orbital spread, which is the square root of the second moment of the orbital, is reported. In this paper, we will refer to the orbital spread as the second moment orbital spread (σ_2) to distinguish it from what we term the fourth moment orbital spread (σ_4), which is the fourth root of the fourth moment of the orbital. σ_4 provides a second essential characterization of the locality of an orbital, which reflects more directly the thickness of the orbital tail than σ_2 does. Thus, to fully characterize the locality of a set of orbitals, we suggest to report σ_2 and σ_4 for the most de-local orbitals (orbitals with the greatest σ_2 and σ_4) in the set.

In Sec. II, we develop the localization function for the fourth moment minimization, and in Sec. III, we derive the equations needed for a minimization of the fourth moment localization function using the trust region method. In Sec. IV, we present results for orbitals obtained using a fourth moment minimization. The orbitals are compared to the orbitals obtained from a second moment minimization both with and without imposed powers, and also with Pipek–Mezey localizations. Section V contains a summary and some concluding remarks.

II. SECOND AND QUARTIC MOMENT LOCALIZATION FUNCTIONS

The locality of an orbital $|p\rangle = \phi_p$ can be associated with its second moment (variance). In one dimension, the second moment of the orbital is defined as

$$\mu_2^p = \langle p|(\hat{x} - \langle p|\hat{x}|p\rangle)^2|p\rangle = \int (x - \bar{x}_p)^2 \phi_p^* \phi_p \, dx, \quad (1)$$

where $\int \phi_p^* \phi_p \, dx = 1$ and $\bar{x}_p = \int x\phi_p^* \phi_p \, dx$. The second moment of an orbital is thus defined in terms of an integration of the density $\phi_p^* \phi_p$ over a second moment weight function $(x - \bar{x}_p)^2$. A local orbital has a small second moment since the bulk of the orbital—represented by its density—will be close to its average position.

For a set of orbitals $\{|p\rangle, |q\rangle \cdots \}$ (e.g., a set of occupied or virtual HF orbitals), the set which on average is most local may be obtained using the Boys localization scheme where the second moment (SM) localization function

$$\xi_{SM} = \sum_p \mu_2^p \quad (2)$$

is minimized carrying out a sequence of orthogonal transformations between the orbitals. When ξ_{SM} is minimized, the orbitals are rotated in order to get their deviation from the mean position $\bar{x}_p$ to be as small as possible. Positions far from $\bar{x}_p$ (where $|x - \bar{x}_p| \gg 0$) are given a large weight and the orbitals therefore are rotated to get a density, which is small in this region. The Boys localization is thus in that sense directly directed toward getting orbitals, where the bulk of the orbital is confined to a small volume in space. The second moment orbital spread of orbital $|p\rangle$
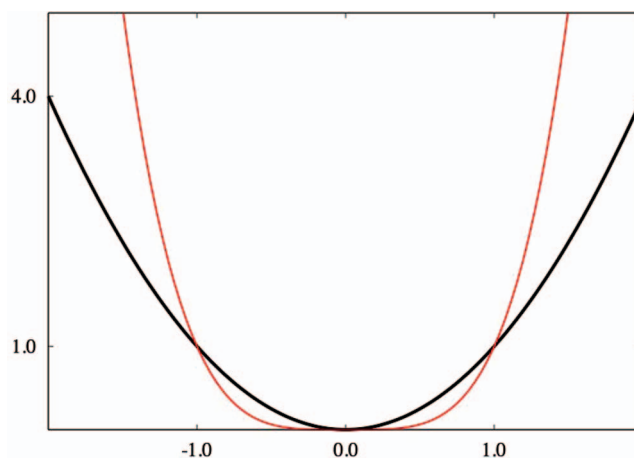
$$\sigma_2^p = \sqrt{\mu_2^p}, \quad (3)$$

FIG. 1. The second (black) and fourth (red) moment weight functions plotted on the interval $[-2,2]$.

which often is denoted just the orbital spread, may be used as a measure for describing where the bulk of the orbital is located.

When minimizing the second moments for a set of orbitals, no particular considerations are given to the shape of the local orbitals, except that the bulk of the density is required to be local. However, an orbital with a small σ_2 may still have a thick tail and such orbitals will be inefficient to use in local correlation methods. To obtain orbitals where thick tails are reduced, we may replace the second moment weight function in Eq. (1) with a fourth moment (FM) weight function.

$$\mu_4^p = \langle p|(\hat{x} - \langle p|\hat{x}|p\rangle)^4|p\rangle \quad (4)$$

and minimize

$$\xi_{FM} = \sum_p \mu_4^p. \quad (5)$$

To understand what is happening in a FM minimization compared to a SM minimization, we have in Fig. 1 plotted the weight function $(\hat{x} - \langle p|\hat{x}|p\rangle)^4$ and $(\hat{x} - \langle p|\hat{x}|p\rangle)^2$. From Fig. 1, it is clear that the weight function is larger in the region $|\hat{x} - \langle p|\hat{x}|p\rangle| > 1$ in a FM minimization than in a SM minimization and opposite for $|\hat{x} - \langle p|\hat{x}|p\rangle| < 1$. In a FM minimization, there is therefore a larger penalty for densities where $|x| > 1$ and a smaller penalty for densities where $|x| < 1$. In Fig. 2, we have for a SM minimization displayed an idealized normal distribution function for the density for an orbital with a second moment orbital spread σ_2. The regions with $|x| \leq \sigma_2$ and $|x| \leq 2\sigma_2$ from the center of the distribution function contain 68% and 95%, respectively, of the total orbital density. An idealized illustration displaying the distribution function obtained from a FM minimization is also plotted in Fig. 2. In the range $|x| < 1$, the FM weight function may pull density away from the center region compared to a SM weight function. For $|x| > 1$, the density is pushed toward the center. Since this range contains a large fraction of the density, a FM minimization may lead to σ_2 values that are smaller than the ones obtained in a SM minimization. The FM minimizations give distributions with much thinner tails than SM minimizations because of the large penalty that is
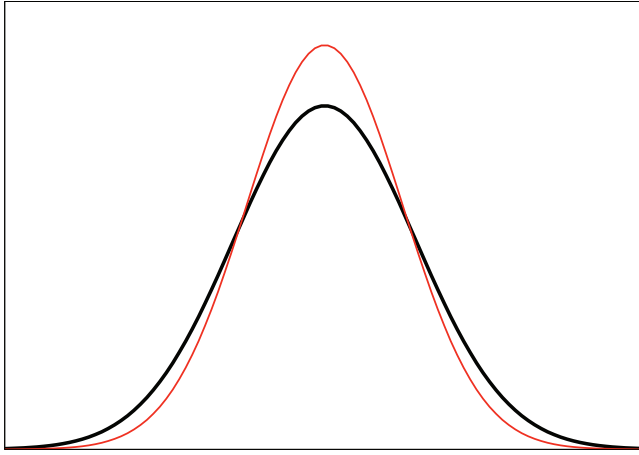
FIG. 2. Illustrations of distribution functions, which ideally would be obtained from a second moment (black) and fourth moment (red) minimization.

imposed by the weight function for $|x| \gg 1$. The fourth moment orbital spread

$$\sigma_4^p = \sqrt[4]{\mu_4^p},\tag{6}$$

thus, measure more directly than σ_2^p how the density decays in the outer and tail region.

In local correlation methods, it is important that the least local orbital is as local as possible, i.e., that there are no outliers. This may be accomplished by introducing a power on the localization function, acting as a penalty for the least local orbitals. Jansik *et al.*[12] introduced powers of the orbital variance in the localization function

$$\xi_{SM}^m = \sum_p \left(\mu_2^p\right)^m,\tag{7}$$

where m is a positive integer. Increasing m increases the penalty for having non-local orbitals. Jansik *et al.* showed that outliers were removed in this way. For fourth moment localization, we similarly consider a localization function that contains powers of the fourth moment

$$\xi_{FM}^m = \sum_p \left(\mu_4^p\right)^m.\tag{8}$$

In three dimensions, the fourth moment may be expressed as

$$\mu_4^p = \langle p|(\mathbf{r} - \bar{\mathbf{r}}_p)^4|p\rangle$$
$$= \int [(\hat{x} - \bar{x}_p)^2 + (\hat{y} - \bar{y}_p)^2 + (\hat{z} - \bar{z}_p)^2]^2 \phi_p^* \phi_p d\tau.\tag{9}$$

In order to characterize the locality of a set of orbitals, the orbital with the largest second moment orbital spread, $|p(\sigma_2^{max})\rangle$,

$$\sigma_2^{max} = \max_p \left[\sigma_2^p\right],\tag{10}$$

and the orbital with the largest fourth moment orbital spread, $|p(\sigma_4^{max})\rangle$,

$$\sigma_4^{max} = \max_p \left[\sigma_4^p\right]\tag{11}$$

must be identified, and both σ_2 and σ_4 will be reported for these orbitals. In general, we recommend to use σ_2 and σ_4 as

measures of locality, regardless of localization scheme since these measures together reflect the locality of both the bulk and tail.

III. MINIMIZATION OF THE FOURTH MOMENT LOCALIZATION FUNCTION USING THE TRUST REGION METHOD

To minimize ξ_{FM}^m, we will use an orthogonal transformation[21] of the orbitals

$$|\tilde{p}\rangle = \exp(-\hat{\kappa})a_{p\sigma}^\dagger|vac\rangle,\tag{12}$$

$$\hat{\kappa} = \sum_{r>s} \kappa_{rs}(E_{rs} - E_{sr}) = \sum_{rs} \kappa_{rs} E_{rs},\tag{13}$$

$$E_{rs} = a_{r\alpha}^\dagger a_{s\alpha} + a_{r\beta}^\dagger a_{s\beta},\tag{14}$$

where $a_{p\sigma}^\dagger$ is a creation operator with σ denoting spin. The summations in Eq. (13) refer to either a set of occupied, or a set of virtual orbitals. The parametrization of ξ_{FM}^m thus becomes

$$\tilde{\xi}_{FM}^m(\boldsymbol{\kappa}) = \sum_p \left(\tilde{\mu}_4^p\right)^m,\tag{15}$$

where for the three dimensional case

$$\tilde{\mu}_4^p = \sum_i \Big[\langle\tilde{p}|\hat{x}_i^4|\tilde{p}\rangle - 4\langle\tilde{p}|\hat{x}_i^3|\tilde{p}\rangle\langle\tilde{p}|\hat{x}_i|\tilde{p}\rangle$$
$$+ 6\langle\tilde{p}|\hat{x}_i^2|\tilde{p}\rangle\langle\tilde{p}|\hat{x}_i|\tilde{p}\rangle^2 - 3\langle\tilde{p}|\hat{x}_i|\tilde{p}\rangle^4\Big]$$
$$+ 2\sum_{i>j} \Big[\langle\tilde{p}|\hat{x}_i^2\hat{x}_j^2|\tilde{p}\rangle - 2(1 + \hat{P}_{ij})\langle\tilde{p}|\hat{x}_i^2\hat{x}_j|\tilde{p}\rangle\langle\tilde{p}|\hat{x}_j|\tilde{p}\rangle$$
$$+ (1 + \hat{P}_{ij})\langle\tilde{p}|\hat{x}_i^2|\tilde{p}\rangle\langle\tilde{p}|\hat{x}_j|\tilde{p}\rangle^2 - 3\langle\tilde{p}|\hat{x}_j|\tilde{p}\rangle^2\langle\tilde{p}|\hat{x}_i|\tilde{p}\rangle^2$$
$$+ 4\langle\tilde{p}|\hat{x}_i\hat{x}_j|\tilde{p}\rangle\langle\tilde{p}|\hat{x}_i|\tilde{p}\rangle\langle\tilde{p}|\hat{x}_j|\tilde{p}\rangle\Big].\tag{16}$$

Subscripts i and j are used to distinguish between the x, y, and z spatial coordinates and the operator $\hat{P}_{ij}$ has been introduced to denote a permutation between indices i and j.

To determine the minimum of $\tilde{\xi}_{FM}^m(\boldsymbol{\kappa})$, a trust region minimization is used. A general discussion of the trust region method may be found in Ref. 22, and a more detailed discussion of the implementation used for orbital localization functions may be found in Ref. 20. The important features are described below.

The trust region method is built around a second-order Taylor series $Q(\boldsymbol{\kappa})$ of $\tilde{\xi}_{FM}^m(\boldsymbol{\kappa})$

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

where $\xi_m^{[0]}$ is the value of $\tilde{\xi}_{FM}^m(\boldsymbol{\kappa})$ at the expansion point ($\boldsymbol{\kappa} = \mathbf{0}$) and $\boldsymbol{\xi}_m^{[1]}$ and $\boldsymbol{\xi}_m^{[2]}$ are the gradient and Hessian of $\tilde{\xi}_{FM}^m(\boldsymbol{\kappa})$, respectively, at the expansion point.

In the trust region method, steps are determined by solving Newton or level-shifted Newton equations using iterative methods. This requires that the gradient $\boldsymbol{\xi}_m^{[1]}$ and the linear transformation of the Hessian on a trial vector $\boldsymbol{\xi}_m^{[2]}\boldsymbol{\kappa}$ are determined. These quantities may be found from a linear expansion

of $\frac{\partial \tilde{\xi}_{FM}^m}{\partial \kappa_{kl}}$.

$$\frac{\partial \tilde{\xi}_{FM}^m}{\partial \kappa_{kl}} = m \sum_p \left(\tilde{\mu}_4^p\right)^{m-1} \frac{\partial \tilde{\mu}_4^p}{\partial \kappa_{kl}}, \qquad (18)$$

where $\tilde{\mu}_4^p$ is given in Eq. (16) and m is a positive integer. The linear expansions of the individual terms in Eq. (18) using $\kappa = 0$ as the expansion point are given by

$$\left(\tilde{\mu}_4^p\right)^{m-1} = f_p^{[0]} + \kappa^T f_P^{[1]} + \mathcal{O}(\kappa^2), \qquad (19)$$

$$\frac{\partial \left(\tilde{\mu}_4^p\right)}{\partial \kappa_{kl}} = \left(g_p^{[0]}\right)_{kl} + \left(\kappa^T g_p^{[1]}\right)_{kl} + \mathcal{O}(\kappa^2). \qquad (20)$$

Using Eq. (16), expressions for $f_p^{[0]}$, $\kappa^T f_P^{[1]}$, $(g_p^{[0]})_{kl}$, and $(\kappa^T g_p^{[1]})_{kl}$ are found. The gradient is not dependent on κ and from the expansions above, we may find the gradient as

$$\left(\xi_m^{[0]}\right)_{kl} = m \sum_p f_p^{[0]} \left(g_p^{[0]}\right)_{kl}. \qquad (21)$$

The linear transformations of the Hessian on a trial vector κ consist of the terms in the expansion that depend linearly on κ and may be found from

$$\left(\xi_m^{[2]} \kappa\right)_{kl} = m \sum_p \left[f_p^{[0]}\left(\kappa^T g_p^{[1]}\right)_{kl} + \left(\kappa^T f_p^{[1]}\right)\left(g_p^{[0]}\right)_{kl}\right]. \qquad (22)$$

The explicit expressions for the gradient and linear transformation of the Hessian on a trial vector are listed in the Appendix.

IV. RESULTS

In this section, we report results from fourth moment orbital localizations for various molecular systems. σ_2 and σ_4 are computed for all orbitals, and the locality of an orbital set is described in terms of σ_2 and σ_4 for the orbital in the set with the largest σ_2, $|p(\sigma_2^{max})\rangle$, and for the orbital with the largest σ_4, $|p(\sigma_4^{max})\rangle$. Arachidic acid ($C_{20}O_2H_{40}$) is used as an example of a one-dimensional system and the I27$_{SS}$ domain ($C_{124}N_{36}O_{37}S_3H_{192}$) of the titin protein as a three-dimensional example. We will for simplicity refer to the I27$_{SS}$ domain of titin as just titin. In all calculations, we will employ Dunning's correlation consistent basis sets cc-pVXZ[23] with cardinal numbers X = D, T, and Q. All starting orbitals used are the least-change orbitals of Ziolkowski et al.,[11] and the localizations were performed using a trust region minimization.[20] No convergence issues were encountered, and for a detailed account on convergence rates for localization functions, we refer to Ref. 20.

In Sec. IV A, we compare the locality measures of orbitals that are obtained in a FM and SM minimization using powers m = 1, 2, and 3, to investigate the effect of m. In Sec. IV B, the shape of the bulk and tail regions in FM and SM minimization are analyzed. For both Secs. IV A and IV B, we focus on examining the locality of the virtual orbitals since these are much more difficult to localize than the occupied orbitals. In Sec. IV C, FM localization with m = 2 as penalty is compared to Boys and Pipek–Mezey localizations for a variety of molecular systems. All orbital

TABLE I. σ_2 and σ_4 values for the virtual orbitals of arachidic acid with σ_2^{max} and σ_4^{max} listed for various basis sets and localization criteria.

		DZ		TZ		QZ	
		σ_2 (a.u.)	σ_4 (a.u.)	σ_2 (a.u.)	σ_4 (a.u.)	σ_2 (a.u.)	σ_4 (a.u.)
ξ_{FM}^1	$|p(\sigma_2^{max})\rangle$	2.70	3.37	2.99	3.75	2.98	4.27
	$|p(\sigma_4^{max})\rangle$	2.35	3.43	2.62	3.84	2.87	4.42
ξ_{FM}^2	$|p(\sigma_2^{max})\rangle$	2.26	2.94	2.18	2.84	2.20	3.02
	$|p(\sigma_4^{max})\rangle$	2.16	3.03	2.09	2.98	2.18	3.02
ξ_{FM}^3	$|p(\sigma_2^{max})\rangle$	2.25	2.92	2.17	2.82	2.12	2.81
	$|p(\sigma_4^{max})\rangle$	2.15	2.95	2.07	2.91	2.11	2.94
ξ_{SM}^1	$|p(\sigma_2^{max})\rangle$	3.19	3.99	3.57	4.55	3.81	5.08
	$|p(\sigma_4^{max})\rangle$	3.19	3.99	3.03	4.59	3.31	5.12
ξ_{SM}^2	$|p(\sigma_2^{max})\rangle$	2.36	3.40	2.43	3.66	2.53	3.49
	$|p(\sigma_4^{max})\rangle$	2.35	3.46	2.42	3.66	2.45	3.66
ξ_{SM}^3	$|p(\sigma_2^{max})\rangle$	2.26	3.22	2.24	3.31	2.14	3.17
	$|p(\sigma_4^{max})\rangle$	2.26	3.27	2.24	3.31	2.03	3.25
PAO	$|p(\sigma_2^{max})\rangle$	3.13	3.40	3.49	3.76	3.82	4.11
	$|p(\sigma_4^{max})\rangle$	2.90	3.79	3.49	3.76	3.82	4.11

visualizations are created using the UCSF Chimera program package.[24]

A. Investigating effects of the power *m*

We will now investigate the effect of an increasing power m on the FM localization function. In Table I, we report the result of a ξ_{FM}^m, m = 1, 2, and 3, minimization of the virtual orbitals for arachidic acid for cardinal numbers X = D, T, and Q. We report the locality measures σ_2 and σ_4 for $|p(\sigma_2^{max})\rangle$ and $|p(\sigma_4^{max})\rangle$. For comparison, we also report locality results from a ξ_{SM}^m, m = 1, 2, and 3 minimization, and the locality of the PAOs.

From Table I, we see that σ_2^{max} and σ_4^{max} changes drastically when going from m = 1 to m = 2, but do not change much when increasing m from 2 to 3. $|p(\sigma_2^{max})\rangle$ and $|p(\sigma_4^{max})\rangle$ have similar σ_2 and σ_4 values. In fact, the distribution of orbitals with σ_2 and σ_4 values in the maximum region is rather dense, as can be seen from the plots in Figs. 3 and 4 where σ_2 and σ_4 are displayed for all virtual orbitals obtained from FM and SM minimizations, respectively, for m = 1 (top), 2 (middle), and 3 (bottom) and using the cc-pVTZ basis set. Figs. 3 and 4 substantiate that it is important to impose a penalty (use m = 2) but increasing this penalty (use m = 3) has only minor effects. This observation is in accordance with the experience from ξ_{SM}^2 localizations,[12] where it was found that m = 2 was a sufficient penalty. From Fig. 3, we also see that for the FM localizations with m = 2, all virtual orbitals have σ_2 values close to 2.0 a.u. and σ_4 values slightly below 3.0 a.u., leaving little space for improving the locality by considering higher than fourth moment localizations.

For titin, results—similar to the ones for arachidic acid— are reported in Table II. The trends for the titin results are the same as for arachidic acid, even though the shifts in locality going from m = 1 to m = 2 are larger for titin than for arachidic acid.
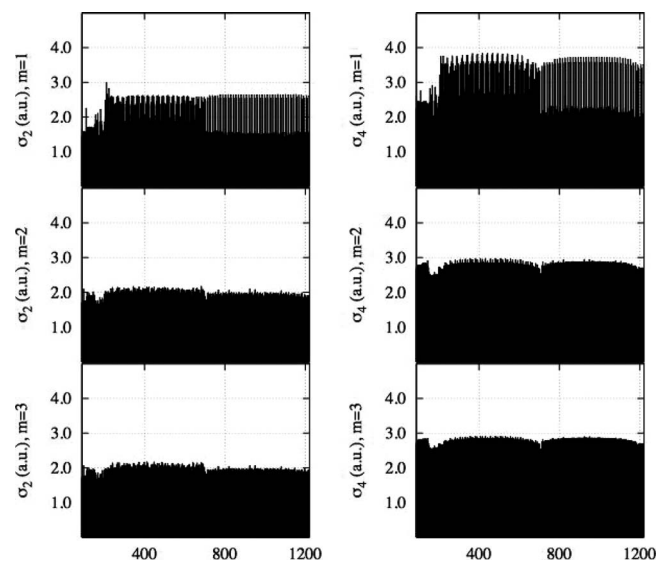
FIG. 3. σ_2 (left) and σ_4 (right) plotted for all virtual orbitals of arachidic acid in a cc-pVTZ basis localized using ξ_{FM}^1 (top), ξ_{FM}^2 (middle), and ξ_{FM}^3 (bottom).

TABLE II. σ_2 and σ_4 values for the virtual orbitals of titin with σ_2^{max} and σ_4^{max} listed for various basis sets and localization criteria.

		DZ		TZ	
		σ_2 (a.u.)	σ_4 (a.u.)	σ_2 (a.u.)	σ_4 (a.u.)
ξ_{FM}^1	$\|p(\sigma_2^{max})\rangle$	3.40	4.06	3.91	5.16
	$\|p(\sigma_4^{max})\rangle$	3.26	4.35	3.73	5.18
ξ_{FM}^2	$\|p(\sigma_2^{max})\rangle$	2.47	3.08	2.45	3.14
	$\|p(\sigma_4^{max})\rangle$	2.42	3.16	2.43	3.21
ξ_{FM}^3	$\|p(\sigma_2^{max})\rangle$	2.46	3.04	2.40	3.10
	$\|p(\sigma_4^{max})\rangle$	2.40	3.10	2.38	3.11
ξ_{SM}^1	$\|p(\sigma_2^{max})\rangle$	3.48	4.21	3.92	5.00
	$\|p(\sigma_4^{max})\rangle$	3.21	4.35	3.62	5.21
ξ_{SM}^2	$\|p(\sigma_2^{max})\rangle$	2.71	3.34	2.76	3.84
	$\|p(\sigma_4^{max})\rangle$	2.60	3.62	2.76	3.84
ξ_{SM}^3	$\|p(\sigma_2^{max})\rangle$	2.52	3.41	2.49	3.33
	$\|p(\sigma_4^{max})\rangle$	2.52	3.41	2.19	3.46
PAO	$\|p(\sigma_2^{max})\rangle$	4.71	5.63	3.68	4.15
	$\|p(\sigma_4^{max})\rangle$	4.65	5.63	2.94	4.50

For the occupied orbitals, we observe trends similar to the ones for the virtual orbitals even though the effects for the occupied orbitals are less pronounced than for the virtual orbitals. As a conclusion, we recommend to use $m = 2$ for FM localizations.

B. The locality of orbitals in fourth moment minimization

In this section, we examine in more detail the orbital locality obtained from a FM minimization with $m = 2$, i.e., using ξ_{FM}^2. For comparison, we present results using ξ_{SM}^2. We start out considering the locality for the virtual orbitals, where we also present the locality of PAOs. Then we examine the
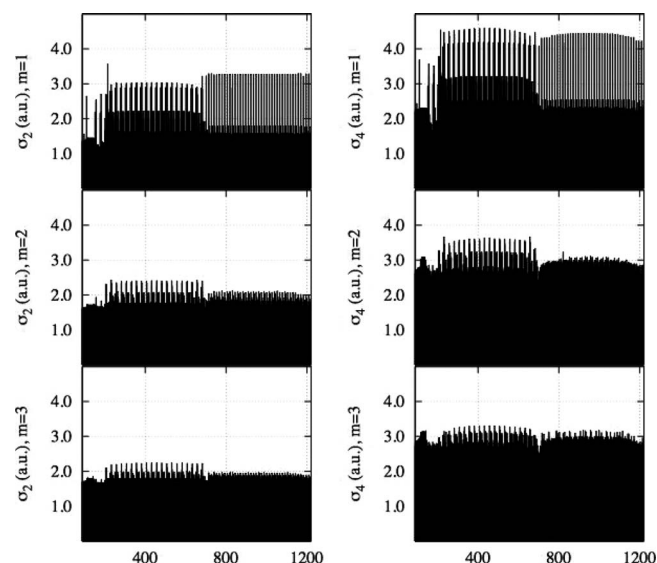
locality of the obtained occupied orbitals. For the occupied orbitals, we only present results for titin, since the arachidic acid occupied orbitals are much less complicated and easily localized.

1. Virtual orbitals

We will now compare the locality of the virtual orbitals obtained in ξ_{FM}^2 and ξ_{SM}^2 localizations for arachidic acid. We first consider the difference in tail decay between FM and SM optimized orbitals. To do this, we have in Fig. 5 for a ξ_{FM}^2 minimization in a cc-pVTZ basis plotted the virtual orbital with the largest fourth moment spread, $|p(\sigma_4^{max})\rangle$, ($\sigma_2 = 2.09$ a.u., $\sigma_4 = 2.98$ a.u.) for the contour values 0.03, 0.003, and 0.001.

FIG. 4. σ_2 (left) and σ_4 (right) plotted for all virtual orbitals of arachidic acid in a cc-pVTZ basis localized using ξ_{SM}^1 (top), ξ_{SM}^2 (middle), and ξ_{SM}^3 (bottom).

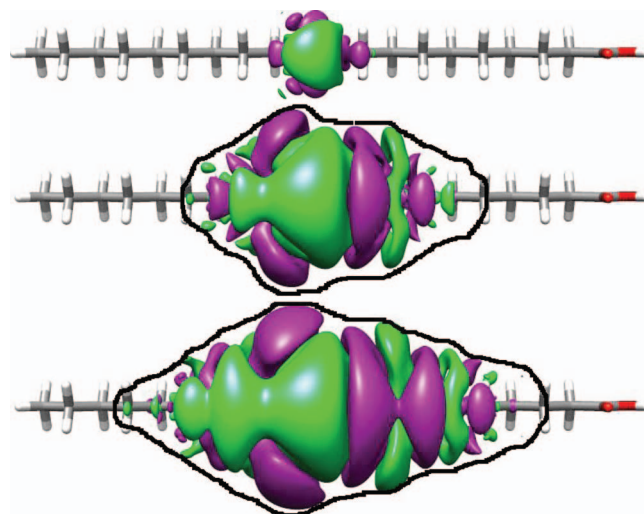FIG. 5. The least local virtual orbital $|p(\sigma_4^{max})\rangle$ ($\sigma_2 = 2.09$ a.u. and $\sigma_4 = 2.98$ a.u.) for arachidic acid in a cc-pVTZ basis obtained from a ξ_{FM}^2 minimization plotted using contour values 0.03 (top), 0.003 (middle), and 0.001 (bottom). A box representing the orbital of Fig. 6 is superimposed to illustrate the more rapid tail decay of the fourth moment optimized orbitals.
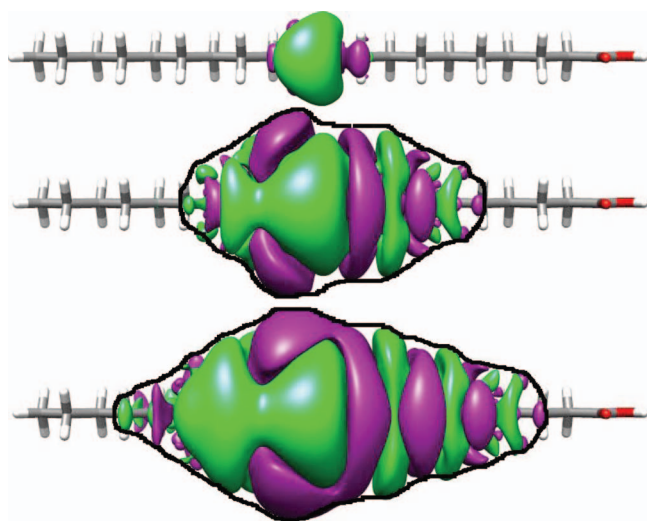
FIG. 6. The virtual orbital centered at the same position as the orbital in Fig. 5 but obtained from a ξ_{SM}^2 localization ($\sigma_2 = 2.41$ a.u. and $\sigma_4 = 3.63$) for arachidic acid in a cc-pVTZ basis, plotted using contour values 0.03 (top), 0.003 (middle), and 0.001 (bottom). A box encasing the orbital is added to use for comparison against the orbital presented in Fig. 5.

For comparison, we have in Fig. 6 plotted the orbital that in a ξ_{SM}^2 minimization is centered on the same carbon atom as the orbital in Fig. 5. This orbital has $\sigma_2 = 2.41$ a.u. and $\sigma_4 = 3.63$ a.u. For the contour values 0.003 and 0.001, we have for the SM optimized orbital (Fig. 6), drawn a black line marking the boundary of the orbital. The boundary is superimposed on the FM optimized orbital in Fig. 5. From Fig. 5, we clearly see that the ξ_{FM}^2 optimized orbital exhibits a more

rapid tail decay than the ξ_{SM}^2 optimized orbital (boundary of ξ_{SM}^2 optimized orbital is denoted by the black line).

To get more insight into the structure of the FM and SM minimized orbitals and their different tail decay, we have in Fig. 7 plotted the distribution functions (left) and the cumulative distribution functions (right) for the orbitals of Figs. 5 and 6 where the integration is performed along the internuclear axis. σ_2 along the nuclear axis is 1.22 a.u. for the FM optimized orbital and 1.49 a.u. for the SM optimized orbital. In Fig. 7 (left), we have also plotted an idealized one-dimensional normal distribution function for the SM optimized orbital ($\sigma_2 = 1.49$ a.u.). From Fig. 7, we see that the idealized distribution function is broader than the distribution function obtained from integration over the SM and FM optimized orbitals. This occurs because the penalty $m = 2$ that is imposed on the orbitals forces the density toward the orbital center both in the SM and the FM localization. For the cumulative distribution functions, the more rapid tail decay is reflected by the fact that the cumulative distribution function for the FM optimized orbital approaches the Heaviside step function more than the SM optimized orbital does, emphasizing that density is being pushed from the tail region toward the center region for the FM optimized orbitals.

Looking at the results in Table I, we see that the FM minimization in general gives lower σ_2^{max} values than a SM minimization. For example, for a TZ basis, FM minimization using $m = 2$ gives $\sigma_2^{max} = 2.18$ a.u., while a SM minimization using $m = 2$ gives $\sigma_2^{max} = 2.43$ a.u. At a first glance, this is a peculiar result since σ_2 (via μ_2^p) is directly minimized in a SM minimization, but not in a FM minimization. The observation is understood from the theoretical analysis in Sec. II, which showed that the bulk of the density obtained
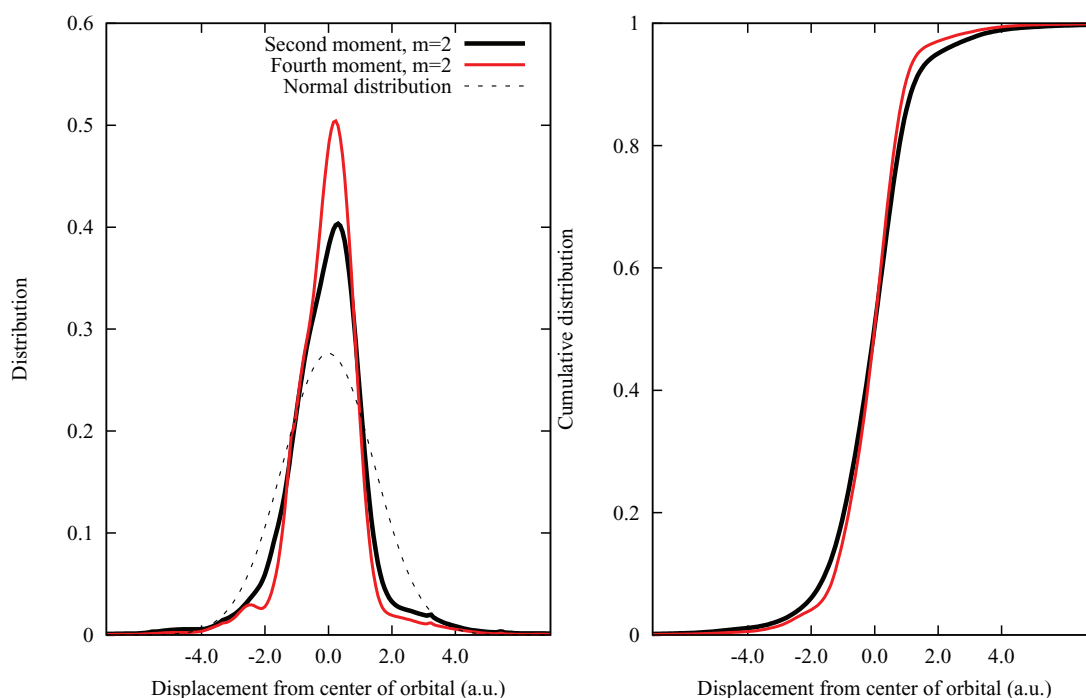

FIG. 7. The distribution functions (left) and cumulative distribution functions (right) for the orbitals plotted in Fig. 5 (red curve) and Fig. 6 (black curve). In addition, a normal distribution (dashed line) with the same orbital spread along the internuclear axis as the orbital represented by the black curve is illustrated.
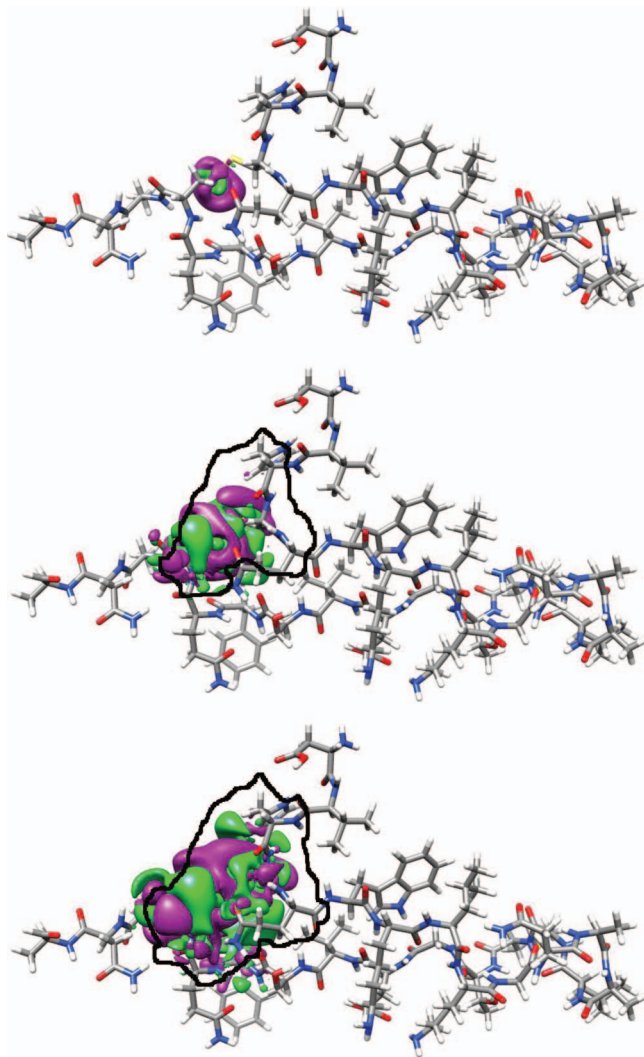
FIG. 8. The most de-local virtual orbital $|p(\sigma_4^{max})\rangle$ ($\sigma_2^{max} = 2.42$ a.u., $\sigma_4^{max} = 3.16$ a.u.) for titin in a cc-pVDZ basis obtained from a ξ_{FM}^2 minimization plotted using contour values 0.03 (top), 0.003 (middle), and 0.001 (bottom). The box which was encasing the orbitals shown in Fig. 9 is superimposed to illustrate the more rapid tail decay of the fourth moment optimized orbitals.

FIG. 9. The most de-local virtual orbital $|p(\sigma_4^{max})\rangle$ ($\sigma_2^{max} = 2.60$ a.u., $\sigma_4^{max} = 3.62$ a.u.) for titin in a cc-pVDZ basis obtained from a ξ_{SM}^2 minimization plotted using contour values 0.03 (top), 0.003 (middle), and 0.001 (bottom). A box encasing the orbital is added to use for comparison against the orbitals presented in Fig. 8.

from a FM minimization will be more compact compared to a SM minimization, because the density is pushed toward the center region in a FM minimization.

We also examine the locality for the virtual orbitals of titin obtained using a ξ_{FM}^2 optimization. From Table II, we see that for the cc-pVDZ basis $\sigma_2^{max} = 2.47$ a.u. and $\sigma_4^{max} = 3.16$ a.u. from a ξ_{FM}^2 minimization, while the corresponding numbers are $\sigma_2^{max} = 2.71$ a.u. and $\sigma_4^{max} = 3.62$ a.u. from a ξ_{SM}^2 minimization. Both the maximum values for σ_2 and σ_4 are thus reduced in a FM minimization compared to a SM minimization. In Figs. 8 and 9, we have plotted the orbitals with σ_4^{max} for a ξ_{FM}^2 and a ξ_{SM}^2 minimization, respectively. The results from Table II and from Figs. 8 and 9 confirm that the trends observed for arachidic acid are also found for titin. For example, the orbital plot in Fig. 8 clearly shows how the tail is reduced in a FM localization compared to a SM localization. Comparing the σ_2^{max} and σ_4^{max} values obtained from a FM minimzation with the locality of the PAOs for titin

(Table II), we see that the PAOs are much less local than FM localized orbitals.

2. Occupied orbitals

We now proceed to examine the locality of the occupied orbitals obtained from a FM minimization. In Table III, we give locality results for the valence orbitals of titin. As seen from Table III, σ_2 and σ_4 are very similar whether a SM or

TABLE III. σ_2 and σ_4 values for the occupied orbitals with σ_2^{max} and σ_4^{max} of titin in various basis sets.

		DZ		TZ	
		σ_2 (a.u.)	σ_4 (a.u.)	σ_2 (a.u.)	σ_4 (a.u.)
ξ_{FM}^2	$|p(\sigma_2^{max})\rangle$	2.20	2.91	2.21	2.72
	$|p(\sigma_4^{max})\rangle$	2.20	2.91	2.20	2.91
ξ_{SM}^2	$|p(\sigma_2^{max})\rangle$	2.19	2.94	2.20	2.96
	$|p(\sigma_4^{max})\rangle$	2.19	2.94	2.19	2.96

TABLE IV. σ_2 and σ_4 values for the virtual orbitals with σ_2^{max} and σ_4^{max} for various molecules.

		ξ_{FM}^2		Boys (ξ_{SM}^1)		Pipek–Mezey	
		σ_2 (a.u.)	σ_4 (a.u.)	σ_2 (a.u.)	σ_4 (a.u.)	σ_2 (a.u.)	σ_4 (a.u.)
Alanine20/cc-pvDZ	$\lvert p(\sigma_2^{max})\rangle$	2.22	2.82	3.32	4.26	3.63	4.33
	$\lvert p(\sigma_4^{max})\rangle$	2.20	3.02	3.32	4.26	3.48	4.40
Alanine20/cc-pvTZ	$\lvert p(\sigma_2^{max})\rangle$	2.17	2.89	3.72	4.74	4.28	5.40
	$\lvert p(\sigma_4^{max})\rangle$	1.99	2.91	3.49	4.94	4.26	5.40
Aβ_{1-42}/cc-pVDZ	$\lvert p(\sigma_2^{max})\rangle$	2.46	3.03	3.53	4.32	4.11	5.48
	$\lvert p(\sigma_4^{max})\rangle$	2.38	3.14	3.32	4.44	4.11	5.48
C_{60}/cc-pVTZ	$\lvert p(\sigma_2^{max})\rangle$	2.27	3.23	3.23	4.46	5.05	6.30
	$\lvert p(\sigma_4^{max})\rangle$	2.25	3.24	2.92	4.55	5.05	6.30
Insulin/cc-pVDZ	$\lvert p(\sigma_2^{max})\rangle$	2.49	3.07	2.58	3.54	4.16	5.57
	$\lvert p(\sigma_4^{max})\rangle$	2.38	3.33	3.35	4.48	4.16	5.57
Graphene/cc-pVDZ	$\lvert p(\sigma_2^{max})\rangle$	2.63	3.58	3.10	4.32	4.55	5.93
	$\lvert p(\sigma_4^{max})\rangle$	2.51	3.65	2.71	4.69	3.66	6.92

a FM localization is carried out. However, when the orbitals are studied in greater detail, we see that tails are reduced if a fourth moment minimization is carried out.

C. Fourth moment localization compared to Boys and Pipek–Mezey localizations

In this section, we compare the locality of orbital sets for various molecular systems, using a ξ_{FM}^2 localization and using Boys and Pipek–Mezey localizations. Boys and Pipek–Mezey localizations have so far been the preferred localization schemes for obtaining local orbitals. The molecules considered are:

- An alpha-helix containing 20 alanine residues, denoted alanine(20), in a cc-pVDZ and cc-pVTZ basis (1924 and 4458 basis functions, respectively).
- The amyloid β-peptide Aβ_{1-42}[25] in a cc-pVDZ basis (6106 basis functions).
- A graphene sheet ($C_{106}H_{28}$) in a cc-pVDZ basis (1624 basis functions).

- C_{60} in a cc-pVTZ basis (1800 basis functions).
- Insulin ($C_{257}N_{65}O_{77}S_6H_{382}^-$) in a cc-pVDZ basis (7604 basis functions).

In Table IV, we report the locality for the virtual orbitals for the molecular systems described above. From Table IV, it is seen that the locality of ξ_{FM}^2 localized virtual orbitals has a rather narrow range for the different molecular systems and also when increasing the cardinal number of the basis sets. The Boys and Pipek–Mezey localized orbitals are seen to be less local than the ξ_{FM}^2 localized orbitals and also showing a larger range for σ_2^{max} and σ_4^{max} values for the different molecular systems. For example, the σ_2^{max} values in Table IV for ξ_{FM}^2 localizations are in the range 2.17–2.63 a.u., while for Boys and Pipek–Mezey localizations the range is 2.58–3.72 a.u. and 3.63–5.05 a.u., respectively. For σ_4^{max}, the range is 2.91–3.65 a.u. for ξ_{FM}^2 localizations, while it is 4.16–4.94 a.u. and 4.40–6.92 a.u. for Boys and Pipek–Mezey localizations, respectively.

In Table V, we report localization results for the occupied orbitals of the molecular systems described above. The

TABLE V. σ_2 and σ_4 values for the occupied orbitals with σ_2^{max} and σ_4^{max} for various molecules.

		ξ_{FM}^2		Boys (ξ_{SM}^1)		Pipek–Mezey	
		σ_2 (a.u.)	σ_4 (a.u.)	σ_2 (a.u.)	σ_4 (a.u.)	σ_2 (a.u.)	σ_4 (a.u.)
Alanine20/cc-pvDZ	$\lvert p(\sigma_2^{max})\rangle$	1.62	2.15	1.70	2.25	1.90	2.43
	$\lvert p(\sigma_4^{max})\rangle$	1.59	2.18	1.70	2.25	1.90	2.43
Alanine20/cc-pvTZ	$\lvert p(\sigma_2^{max})\rangle$	1.61	2.15	1.72	2.29	1.90	2.44
	$\lvert p(\sigma_4^{max})\rangle$	1.58	2.16	1.72	2.29	1.90	2.44
Aβ_{1-42}/cc-pVDZ	$\lvert p(\sigma_2^{max})\rangle$	2.16	2.65	2.12	2.86	2.73	3.24
	$\lvert p(\sigma_4^{max})\rangle$	2.14	2.77	2.12	2.86	2.72	3.25
C_{60}/cc-pVTZ	$\lvert p(\sigma_2^{max})\rangle$	2.22	3.24	2.36	3.45	2.91	3.91
	$\lvert p(\sigma_4^{max})\rangle$	2.22	3.24	2.36	3.45	2.91	3.91
Insulin/cc-pVDZ	$\lvert p(\sigma_2^{max})\rangle$	2.17	2.67	2.17	2.89	2.76	3.25
	$\lvert p(\sigma_4^{max})\rangle$	2.15	2.78	2.17	2.89	2.72	3.26
Graphene/cc-pVDZ	$\lvert p(\sigma_2^{max})\rangle$	2.58	4.15	3.37	5.60	3.83	7.25
	$\lvert p(\sigma_4^{max})\rangle$	2.54	4.17	3.37	5.60	3.83	7.25

results are presented as for the virtual orbitals reported in Table IV. The results show that a ξ_{FM}^2 localization yields the most local occupied orbitals. When comparing the three schemes, the differences between the locality for the occupied orbitals are generally not as big as for the virtual orbitals. However, it is important to note that for some systems (e.g., graphene), the Boys and Pipek–Mezey localization schemes provide very poor locality for the occupied orbitals. Thus, when using these schemes it is extremely important to check the locality of the obtained orbitals, since they may yield only semi-local orbitals for some molecular systems.

V. SUMMARY AND CONCLUDING REMARKS

We have introduced a localization function for occupied and virtual orbitals based on powers of the orbital fourth moment. The power is introduced to act as a penalty on the least local orbitals, i.e., to ensure that no orbitals in the set are significantly less local than the others. It is shown that the orbitals obtained from a fourth moment minimization with power m have a more rapid tail decay than orbitals obtained from a second moment (Boys) minimization with the same power m. Further, the improvement in tail decay is much more pronounced for the virtual than for the occupied orbitals. A penalty value $m = 2$ was shown to be sufficient to obtain a set of orbitals in which all orbitals are properly local. A FM minimization with a penalty $m = 2$ thus may be recommended for obtaining a set of local HF orbitals for local correlation

methods. Further, we propose that for any localization scheme, the locality of the orbitals should be characterized by the second and fourth moment orbital spreads of the least local orbitals in the set.

APPENDIX: EXPRESSIONS FOR GRADIENT AND LINEAR TRANSFORMATION OF HESSIAN ON TRIAL VECTOR

Expanding Eq. (18), inserting the results into Eqs. (21) and (22), and introducing the odd component of a matrix A as

$$\{A\}^o = A - A^T, \tag{A1}$$

and defining $\mathrm{dia}[v]$ as the operation for taking a vector v and transforming it to a diagonal matrix D_v containing the vector elements on the diagonal

$$\mathrm{dia}[v] = D_v, \tag{A2}$$

we arrive at the following expressions for the gradient $\xi_m^{[1]}$ and linear transformation $\xi_m^{[2]}\kappa$:

$$\xi_m^{[1]} = m \sum_i \Big[-2\{x_i^4 \mathrm{dia}[(\mu_4)^{m-1}]\}^o$$

$$+ 8\{x_i^3 \mathrm{dia}[(\mu_4)^{m-1} x_i]\}^o - 12\{x_i^2 \mathrm{dia}[(\mu_4)^{m-1}[x_i]^2]\}^o$$

$$+ 8\{x_i \mathrm{dia}[(\mu_4)^{m-1}[x_i^3]]\}^o - 24\{x_i \mathrm{dia}[(\mu_4)^{m-1}[x_i^2]x_i]\}^o$$

$$+ 24\{x_i \mathrm{dia}[(\mu_4)^{m-1}[x_i]^3]\}^o \Big]$$

$$+ 2 \sum_{i>j} \Big[-\{x_i^2 x_j^2 \mathrm{dia}[(\mu_4)^{m-1}]\}^o$$

$$- 8\{x_i x_j \mathrm{dia}[(\mu_4)^{m-1} x_i x_j]\}^o$$

$$- 4(1 + P_{ij})\{x_j \mathrm{dia}[(\mu_4)^{m-1}[x_i^2]x_j]\}^o$$

$$+ 4(1 + P_{ij})\{x_j \mathrm{dia}[(\mu_4)^{m-1}[x_i^2 x_j]]\}^o$$

$$+ 12(1 + P_{ij})\{x_j \mathrm{dia}[(\mu_4)^{m-1}[x_i]^2 x_j]\}^o$$

$$- 8(1 + P_{ij})\{x_j \mathrm{dia}[(\mu_4)^{m-1} x_i [x_i x_j]]\}^o$$

$$- 2(1 + P_{ij})\{x_i^2 \mathrm{dia}[(\mu_4)^{m-1}[x_j]^2]\}^o$$

$$+ 4(1 + P_{ij})\{x_j^2 x_i \mathrm{dia}[(\mu_4)^{m-1} x_i]\}^o \Big], \tag{A3}$$

$$\boldsymbol{\xi}_m^{[2]}\boldsymbol{\kappa} = m \sum_{x=x,y,z} \big[-2\{(\boldsymbol{\kappa}\boldsymbol{x}^4 - \boldsymbol{x}^4\boldsymbol{\kappa})\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}]\}^o$$

$$+ 8\{\boldsymbol{x}^3\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}(\boldsymbol{\kappa}\boldsymbol{x} - \boldsymbol{x}\boldsymbol{\kappa})]\}^o + 8\{(\boldsymbol{\kappa}\boldsymbol{x}^3 - \boldsymbol{x}^3\boldsymbol{\kappa})\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}\boldsymbol{x}]\}^o$$

$$+ 8\{\boldsymbol{x}\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}(\boldsymbol{\kappa}\boldsymbol{x}^3 - \boldsymbol{x}^3\boldsymbol{\kappa})\}^o - 24\{\boldsymbol{x}\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}(\boldsymbol{\kappa}\boldsymbol{x}^2 - \boldsymbol{x}^2\boldsymbol{\kappa})\boldsymbol{x}\}^o$$

$$- 24\{\boldsymbol{x}\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}[\boldsymbol{x}^2](\boldsymbol{\kappa}\boldsymbol{x} - \boldsymbol{x}\boldsymbol{\kappa})\}^o + 72\{\boldsymbol{x}\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}[\boldsymbol{x}]^2(\boldsymbol{\kappa}\boldsymbol{x} - \boldsymbol{x}\boldsymbol{\kappa})]\}^o$$

$$+ 8\{(\boldsymbol{\kappa}\boldsymbol{x} - \boldsymbol{x}\boldsymbol{\kappa})\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}[\boldsymbol{x}^3]\}^o - 24\{(\boldsymbol{\kappa}\boldsymbol{x} - \boldsymbol{x}\boldsymbol{\kappa})\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}\boldsymbol{x}[\boldsymbol{x}^2]\}^o$$

$$+ 24(\boldsymbol{\kappa}\boldsymbol{x} - \boldsymbol{x}\boldsymbol{\kappa})\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}[\boldsymbol{x}]^3]\}^o - 24\{\boldsymbol{x}^2\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}\boldsymbol{x}(\boldsymbol{\kappa}\boldsymbol{x} - \boldsymbol{x}\boldsymbol{\kappa})]\}^o$$

$$- 12\{(\boldsymbol{\kappa}\boldsymbol{x}^2 - \boldsymbol{x}^2\boldsymbol{\kappa})\mathrm{dia}[(\boldsymbol{\mu}_4)^{m-1}[\boldsymbol{x}]^2]\}^o$$

$$- 2\{\boldsymbol{x}^4\mathrm{dia}[\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]}]\}^o + 8\{\boldsymbol{x}^3\mathrm{dia}[(\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]})\boldsymbol{x}]\}^o$$

$$- 12\{\boldsymbol{x}^2\mathrm{dia}[(\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]})[\boldsymbol{x}]^2]\}^o + 8\{\boldsymbol{x}\mathrm{dia}[(\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]})[\boldsymbol{x}^3]]\}^o$$

$$- 24\{\boldsymbol{x}\mathrm{dia}[(\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]})[\boldsymbol{x}^2]\boldsymbol{x}]\}^o + 24\{\boldsymbol{x}\mathrm{dia}[(\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]})[\boldsymbol{x}]^3]\}^o \big]$$

$$+ 2m \sum_{i>j} \big[-2\{(\boldsymbol{\kappa}\boldsymbol{x}_i^2\boldsymbol{x}_j^2 - \boldsymbol{x}_i^2\boldsymbol{x}_j^2\boldsymbol{\kappa})\mathrm{dia}[\boldsymbol{f}^{[0]}\}^o$$

$$- 8\{(\boldsymbol{\kappa}\boldsymbol{x}_i\boldsymbol{x}_j - \boldsymbol{x}_i\boldsymbol{x}_j\boldsymbol{\kappa})\mathrm{dia}[\boldsymbol{f}^{[0]}\boldsymbol{x}_i\boldsymbol{x}_j]\}^o + (1 + P_{ij})\{\boldsymbol{x}_j\mathrm{dia}[\boldsymbol{f}^{[0]}[\boldsymbol{x}_i^2](\boldsymbol{\kappa}\boldsymbol{x}_j - \boldsymbol{x}_j\boldsymbol{\kappa})]\}^o$$

$$- 4(1 + P_{ij})\{\boldsymbol{x}_j\mathrm{dia}[\boldsymbol{f}^{[0]}(\boldsymbol{\kappa}\boldsymbol{x}_i^2 - \boldsymbol{x}_i^2\boldsymbol{\kappa})\boldsymbol{x}_j]\}^o + 4(1 + P_{ij})\{\boldsymbol{x}_j\mathrm{dia}[\boldsymbol{f}^{[0]}(\boldsymbol{\kappa}\boldsymbol{x}_i^2\boldsymbol{x}_j - \boldsymbol{x}_i^2\boldsymbol{x}_j\boldsymbol{\kappa})]\}^o$$

$$+ 24(1 + P_{ij})\{\boldsymbol{x}_j\mathrm{dia}[\boldsymbol{f}^{[0]}[\boldsymbol{x}_j][\boldsymbol{x}_i](\boldsymbol{\kappa}\boldsymbol{x}_i - \boldsymbol{x}_i\boldsymbol{\kappa})]\}^o - 8(1 + P_{ij})\{\boldsymbol{x}_j\mathrm{dia}[\boldsymbol{f}^{[0]}\boldsymbol{x}_i(\boldsymbol{\kappa}\boldsymbol{x}_i\boldsymbol{x}_j - \boldsymbol{x}_i\boldsymbol{x}_j\boldsymbol{\kappa})]\}^o$$

$$- 8(1 + P_{ij})\{\boldsymbol{x}_j\mathrm{dia}[\boldsymbol{f}^{[0]}\boldsymbol{x}_i(\boldsymbol{\kappa}\boldsymbol{x}_i - \boldsymbol{x}_i\boldsymbol{\kappa})[\boldsymbol{x}_i\boldsymbol{x}_j]]\}^o - 4(1 + P_{ij})\{(\boldsymbol{\kappa}\boldsymbol{x}_j - \boldsymbol{x}_j\boldsymbol{\kappa})\mathrm{dia}[\boldsymbol{f}^{[0]}[\boldsymbol{x}_i^2]\boldsymbol{x}_j]\}^o$$

$$+ 4(1 + P_{ij})\{(\boldsymbol{\kappa}\boldsymbol{x}_j - \boldsymbol{x}_j\boldsymbol{\kappa})\mathrm{dia}[\boldsymbol{f}^{[0]}[\boldsymbol{x}_i^2\boldsymbol{x}_j]\}^o + 12(1 + P_{ij})\{(\boldsymbol{\kappa}\boldsymbol{x}_j - \boldsymbol{x}_j\boldsymbol{\kappa})\mathrm{dia}[\boldsymbol{f}^{[0]}[\boldsymbol{x}_i]^2\boldsymbol{x}_j]\}^o$$

$$- 8(1 + P_{ij})\{(\boldsymbol{\kappa}\boldsymbol{x}_j - \boldsymbol{x}_j\boldsymbol{\kappa})\mathrm{dia}[\boldsymbol{f}^{[0]}[\boldsymbol{x}_i\boldsymbol{x}_j]\boldsymbol{x}_i]\}^o - 4(1 + P_{ij})\{\boldsymbol{x}_i^2\mathrm{dia}[\boldsymbol{f}^{[0]}(\boldsymbol{\kappa}\boldsymbol{x}_j - \boldsymbol{x}_j\boldsymbol{\kappa})\boldsymbol{x}_j]\}^o$$

$$- 2(1 + P_{ij})\{(\boldsymbol{\kappa}\boldsymbol{x}_j^2 - \boldsymbol{x}_j^2\boldsymbol{\kappa})\mathrm{dia}[\boldsymbol{f}^{[0]}[\boldsymbol{x}_i]^2\}^o + 4(1 + P_{ij})\{(\boldsymbol{\kappa}\boldsymbol{x}_i^2\boldsymbol{x}_j - \boldsymbol{x}_i^2\boldsymbol{x}_j\boldsymbol{\kappa})\mathrm{dia}[\boldsymbol{f}^{[0]}\boldsymbol{x}_j]\}^o$$

$$+ 4(1 + P_{ij})\{\boldsymbol{x}_i^2\boldsymbol{x}_j\mathrm{dia}[\boldsymbol{f}^{[0]}(\boldsymbol{\kappa}\boldsymbol{x}_j - \boldsymbol{x}_j\boldsymbol{\kappa})]\}^o - 8(1 + P_{ij})\{\boldsymbol{x}_i\boldsymbol{x}_j\mathrm{dia}[\boldsymbol{f}^{[0]}(\boldsymbol{\kappa}\boldsymbol{x}_i - \boldsymbol{x}_i\boldsymbol{\kappa})\boldsymbol{x}_j]\}^o$$

$$+ 2m \sum_{i>j} \big[-2\{\boldsymbol{x}_i^2\boldsymbol{x}_j^2\mathrm{dia}[\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]}]\}^o - 8\{\boldsymbol{x}_i\boldsymbol{x}_j\mathrm{dia}[(\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]})\boldsymbol{x}_i\boldsymbol{x}_j]\}^o$$

$$- 4(1 + P_{ij})\{\boldsymbol{x}_j\mathrm{dia}[(\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]})[\boldsymbol{x}_i^2]\boldsymbol{x}_j]\}^o + 12(1 + P_{ij})\{\boldsymbol{x}_j\mathrm{dia}[(\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]})[\boldsymbol{x}_i]^2\boldsymbol{x}_j]\}^o$$

$$4(1 + P_{ij})\{\boldsymbol{x}_j\mathrm{dia}[(\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]})[\boldsymbol{x}_i^2\boldsymbol{x}_j]]\}^o - 2(1 + P_{ij})\{\boldsymbol{x}_i^2\mathrm{dia}[(\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]})[\boldsymbol{x}_j]^2]\}^o$$

$$+ 4(1 + P_{ij})\{\boldsymbol{x}_i^2\boldsymbol{x}_j\mathrm{dia}[(\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]})\boldsymbol{x}_j]\}^o \big]$$

$$- \frac{1}{2}\{\boldsymbol{\kappa}\boldsymbol{\xi}_m^{[0]}\}^o. \tag{A4}$$

Subscripts i and j are used to distinguish between the x, y, and z spatial coordinates and the operator $\hat{P}_{ij}$ has been introduced to denote a permutation between indices i and j. The cartesian moment matrix $\boldsymbol{x}_i^n$ is defined by

$$\boldsymbol{x}_i^n = \sum_{p,q} \langle p|x_i^n|q\rangle. \tag{A5}$$

In Table VI, we have specified the dimensions of the quantities entering Eqs. (A3) and (A4).

For the gradient (Eq. (A3)), there are no matrix-matrix multiplications, only vector-matrix operations. The linear

TABLE VI. The dimension of the quantities entering the gradient and the linear transformation of the Hessian on a trial vector. N_{orb} is the number of orbitals in the set. Note: all vectors are defined to be row vectors.

Quantity	Dimensions
$\boldsymbol{\xi}_m^{[1]}$	$N_{orb} \times N_{orb}$
$\boldsymbol{\xi}_m^{[2]}\boldsymbol{\kappa}$	$N_{orb} \times N_{orb}$
$\boldsymbol{\kappa}$	$N_{orb} \times N_{orb}$
$\boldsymbol{x}_i^n$	$N_{orb} \times N_{orb}$
$\boldsymbol{\kappa}\boldsymbol{x}_i^n$	$N_{orb} \times N_{orb}$
$(\boldsymbol{\mu}_4)^{m-1}$	$1 \times N_{orb}$
$\boldsymbol{f}^{[0]}$	$1 \times N_{orb}$
$\boldsymbol{\kappa}^T\boldsymbol{f}^{[1]}$	$1 \times N_{orb}$

transformation of the Hessian on a trial vector (Eq. (A4)) contains 25 matrix-matrix multiplications. One multiplication is the gradient multiplied with the trial vector and the remaining 24 are multiplications of each of the cartesian moment matrices with the trial vector.

All matrix operations are implemented in parallel and distributed fashion using the block cyclic distribution of data and the ScaLAPACK library.[26] Interprocess communication is done via the message passing interface (MPI) of the MPI library.[27] This allows us to utilize the combined processor power and memory of several computational nodes in demanding evaluations of the linear transformation. The parallelization enables fourth moment minimization to be applied to large molecules with reasonably sized basis sets.

[1] F. Jensen, *Introduction to Computational Chemistry*, 2nd ed. (Wiley, 2007).
[2] J. Pipek, Int. J. Quantum Chem. **36**, 487 (1989).
[3] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).
[4] S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).
[5] S. F. Boys, *Quantum Theory of Atoms, Molecules and Solid State* (Academic, New York, 1966), p. 253.
[6] C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).
[7] W. von Niessen, J. Chem. Phys. **56**, 4290 (1972).
[8] J. E. Subotnik, A. D. Dutoi, and M. Head-Gordon, J. Chem. Phys. **123**, 114108 (2005).
[9] F. Aquilante, T. Pedersen, A. S. de Meras, and H. Koch, J. Chem. Phys. **125**, 174101 (2006).
[10] V. Weber and J. Hutter, J. Chem. Phys. **128**, 064107 (2008).
[11] M. Ziółkowski, B. Jansík, P. Jørgensen, and J. Olsen, J. Chem. Phys. **131**, 124112 (2009).
[12] B. Jansík, S. Høst, K. Kristensen, and P. Jørgensen, J. Chem. Phys. **134**, 194104 (2011).
[13] P. de Silva, M. Giebultowski, and J. Korchowiec, Phys. Chem. Chem. Phys. **14**, 546 (2012).
[14] J. E. Lennard-Jones and J. A. Pople, Proc. R. Soc. London **202**, 166 (1950).
[15] G. Berghold, C. J. Mundy, A. H. Romero, J. Hutter, and M. Parrinello, Phys. Rev. B **61**, 10040 (2000).
[16] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 35109 (2002).
[17] S. Saebø and P. Pulay, Ann. Rev. Phys. Chem. **44**, 213 (1993).
[18] M. Schütz, G. Hetzer, and H.-J. Werner, J. Chem. Phys. **111**, 5691 (1999).
[19] M. Schütz and H.-J. Werner, J. Chem. Phys. **114**, 661 (2001).
[20] I.-M. Høyvik, B. Jansik, and P. Jørgensen, J. Chem. Theory Comput. **8**, 3137 (2012).
[21] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory*, 1st ed. (Wiley, 2000).
[22] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. (Wiley, Chichester, 1987).
[23] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[24] E. Pettersen, T. Goddard, C. Huang, G. Couch, D. Greenblatt, E. Meng, and T. Ferrin, J. Comput. Chem. **13**, 1605 (2004).
[25] B. Strodel, J. W. L. Lee, C. S. Whittleston, and D. J. Wales, J. Am. Chem. Soc. **132**, 13300 (2010).
[26] L. S. Blackford, J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet *et al.*, *ScaLAPACK Users' Guide* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997).
[27] E. Gabriel, G. E. Fagg, G. Bosilca, T. Angskun, J. J. Dongarra, J. M. Squyres, V. Sahay, P. Kambadur, B. Barrett, A. Lumsdaine *et al.*, in *Proceedings of the 11th European PVM/MPI Users' Group Meeting on Open MPI: Goals, Concept, and Design of a Next Generation MPI Implementation* (Euro PVM/MPI, Budapest, Hungary, 2004).