Theor Chem Acc (2005) 114: 276–282
DOI 10.1007/s00214-005-0685-5

REGULAR ARTICLE

Denis Usvyat · Martin Schütz

Nonorthogonal ultralocalized functions and fitted Wannier functions for local electron correlation methods for solids

Received: 20 June 2005 / Accepted: 22 June 2005 / Published online: 28 September 2005

Abstract Despite the formal exponential decay behavior of Wannier functions (WFs), their spatial extent, which is a key parameter determining the computational cost of local correlation calculations for solids, is still rather large. The problems with the localization of the WFs can partly be attributed to their mutual orthogonality. Possibilities of reduction of the spatial extent of the WFs without losing the accuracy of the calculations are investigated. A method for generation of nonorthogonal ultralocalized functions based on maximization of their Löwdin populations is developed. A scheme for fitting of the WFs and nonorthogonal localized functions with a limited support is proposed. The calculations show that by combining both techniques one can obtain quite compact linearly independent localized functions, which may significantly decrease the computational cost in post-HF calculations.

Keywords Wanniee functions · Nonorthogonal localized functions · Local correlation methods · Symmetry in crystals

1 Introduction

Post-HF theories and other computational techniques for electronic structure calculations during many decades have been utilizing the canonical one-electron orbitals. However, these orbitals have the inconvenient property of being delocalized over the whole spatial extent of the system, which leads to high scaling of the computational cost with respect to the system size. On the other hand, it is not essential whether to use the canonical one-electron functions or their linear combinations to form the determinants for the many-electron wavefunctions. Recently, low-order scaling algorithms have been developed for electronic structure calculations, which use localized or Wannier – in case of periodic systems – functions [11,22,25].

Wannier functions (WFs) were initially introduced as far back as 1937 [31] and ever since then they play an important role in the solid state theory. However, for quite a long time their application was restricted mainly to pure theoretical modeling without actual calculations. But in the last decade, after a number of effective methods for numerical generation of WFs had been developed [16,17,29,30,34], they have also been very intensively used in practical calculations of various properties of solids [2,9,13,19,22,24,27,35].

Though WFs are spatially localized, the localization region can be relatively large. As an illustration one can consider the case of diamond. In the linear combination of atomic orbitals (LCAO) approximation, the WFs are represented by the AO expansion coefficients. The coefficients of the four translationally inequivalent WFs (which correspond to the four upper valence bands of the HF reference) larger (in absolute value) than a threshold of 10^{-4} comprise the atomic functions centered in as much as 435 (!) unit cells. Such a wide extent of WFs makes accurate calculations for solids computationally difficult.

The slow decay of WFs is often related to their mutual orthogonality, which gives rise to long range oscillating tails, containing no physically relevant information. For quite sometime now, the possibility to avoid these tails by sacrificing orthogonality has been discussed. Many important results concerning the localization properties of WFs and nonorthogonal localized orbitals (NOLO) (following most of the authors, we don't use the term "Wannier functions" for nonorthogonal localized functions) were obtained for the one-dimensional (1D) case. In 1959, W. Kohn, in his classical work on analytical properties of Bloch and Wannier functions for 1D periodic potentials with a center of inversion, proved the existence of real, symmetry adapted exponentially decaying WFs [14]. In 1961, Andersen showed that in the 1D case, the spread of the NOLOs can be lower than that of the WFs [1]. He introduced the term "ultralocalized" for such nonorthogonal functions. Much later, expansions of such functions with minimal spread in a basis of Gaussians situated on a 1D grid were studied [33]. Recently, He and Vanderbilt [12] showed that the actual decay rate of the 1D Wannier functions $W(r)$ is not just an exponential:

$$W(r) \sim \exp(-hr), \tag{1}$$

D. Usvyat (✉) · M. Schütz
Institut für Physikalische und Theoretische Chemie, Universität Regensburg, Universitätsstraße 31, D-93040 Regensburg, Germany
E-mail: {denis.usvyat, martin.schnetz}@chemie.uni-regensburg.de

but rather an exponential times a power-law prefactor:

$$W(r) \sim r^{-\alpha}\exp(-hr), \tag{2}$$

where h is a parameter depending on the band gap widths, and α is a constant which in 1D takes the values of $3/4$ for WFs, and $1/2$, $3/2$ and $5/2$ for nonorthogonal localized functions of different types. This, at the first glance inessential prefactor of the exponential is actually quite important, since, writing (2) in the logarithmic scale:

$$\ln(W(r)) \sim -\alpha\ln(r) - hr, \tag{3}$$

and taking the actual values of α, reveals the fact that nonorthogonal functions can be essentially more localized than WFs, as the NOLO with $\alpha = 5/2$ can be (for a certain r) some orders of magnitude closer to zero than the WF with $\alpha = 3/4$.

Following the results of the 1D investigations of WFs and nonorthogonal functions, much attention has been paid to the three-dimensional (3D) case, as well (for which analytical results are not always available). Unfortunately, it turned out to be not that straight-forward to construct 3D nonorthogonal functions with a better localization than that of WFs. Surprisingly, real nonorthogonal functions $W_i^{(\mathrm{no})}(\mathbf{r})$ (index i numbers the functions in the reference unit cell) that minimize the spread functional

$$I_i^{(\mathrm{spr})} = \int r^2 (W_i^{(\mathrm{no})})^2 d\mathbf{r} - [\int \mathbf{r}(W_i^{(\mathrm{no})})^2 d\mathbf{r}]^2 \tag{4}$$

do not show any significant difference from their orthogonal counterparts [21]. The same effect was observed when functionals of the following form

$$I_i^{(w_i)} = \int w_i(\mathbf{r})(W_i^{(\mathrm{no})})^2 d\mathbf{r}, \tag{5}$$

where $w_i(\mathbf{r})$ is a compact Gauss- or a delta-function centered in the vicinity of the center of the function $W_i^{(\mathrm{no})}(\mathbf{r})$, were used [8,29]. In Mayer et al. [17] a maximization of LCAO Mulliken net populations of the nonorthogonal functions in the reference cell was proposed. However, since the Mulliken population of a function on a set of atoms can exceed unity due to nonorthogonality of the AO basis, such a procedure still leads to nonorthogonal localized functions with noticeable "tails". Nonorthogonal orbitals have also been constructed by projecting some trial localized functions onto the occupied space [20,21]. The resulting functions appeared to be either again quite similar to the corresponding WFs or, in some cases, even more delocalized than the latter.

An efficient method for generating NOLOs for subsequent quantum Monte-Carlo calculations has been proposed in Reboredo and Williamson [23]. As the calculations have shown, it allows one to obtain very well localized nonorthogonal functions, which drop practically to zero beyond a certain region, for nonconducting crystals and even for metals. The method implies maximizing the functional of the form (5) with the weight function

$$w_i(\mathbf{r}) = \Theta(R_i^{(\mathrm{cut})} - |\mathbf{r} - \mathbf{r}_i|), \tag{6}$$

where $\Theta(x)$ is the Heaviside step-function, $\mathbf{r}_i$ is the centroid of the function $W_i^{(\mathrm{no})}(\mathbf{r})$, and $R_i^{(\mathrm{cut})}$ is the cutoff radius. Thus, the square of the function $W_i^{(\mathrm{no})}(\mathbf{r})$ is maximized inside a sphere of a radius $R_i^{(\mathrm{cut})}$. The process of NOLO construction is preceeded by setting the tolerance for the value $1 - I_i^{(w_i)}$ with w_i defined in (6). Then, one successively increases the cutoff radius, until the chosen tolerance is satisfied. Finally, the $R_i^{(\mathrm{cut})}$ so obtained is used for constructing the NOLOs. The calculations have shown that the method is efficient for generating essentially compact NOLOs, provided the chosen tolerance is sufficiently low. For example, in the case of bulk Silicon, the NOLOs are noticeably more compact than the WFs and tend to concentrate entirely inside the sphere and thus feature no long-range "tails", provided the tolerance is chosen less than 10^{-2} (the cutoff radius then turns out to be larger than 3 a.u.). There are similarities between this method and order-N DFT schemes [28,11], where the nonorthogonal localized orbitals are assumed to be nonzero inside a pre-specified localization region only.

Though the method from Reboredo and Williamson [23] seems to be quite promising, it is not actually applicable for the LCAO calculations due to difficulties with evaluating the integrals of Gaussians over a finite region. But the possibility to get rid of the WFs' tails may have a significant impact on the efficiency of local correlation methods like local Møller-Plesset second order perturbation theory (LMP2). For these methods the most expensive step is the evaluation and transformation of four-index two-electron integrals, where the computational cost is governed by the spatial extent of the AO support of the individual WFs. In this paper, we present techniques for constructing compact WFs or NOLOs in AO basis sets. The intention is to improve the efficiency of the local correlation methods in nonconducting crystals [22]. Nevertheless, the results are general and can also be applied to other cases where the LCAO expansion of WFs is a critical issue. The first results on using these functions in the context of LMP2 are reported.

2 Nonorthogonal ultralocalized functions

The first step in the construction of NOLOs is the specification of an appropriate functional to be minimized or maximized. The LCAO expansion of the WF $W_i(\mathbf{r})$ is defined as

$$W_i(\mathbf{r}) = \sum_{\mu,\mathbf{R}} C_{\mu\mathbf{R},i}\phi_\mu(\mathbf{r} - \mathbf{R}) \equiv \sum_{\boldsymbol{\mu}} C_{\boldsymbol{\mu},i}\phi_{\boldsymbol{\mu}}(\mathbf{r}), \tag{7}$$

where $\phi_\mu(\mathbf{r} - \mathbf{R})$ represents the atomic function centered in the $\mathbf{R}$th unit cell, $C_{\mu\mathbf{R},i}$ is the corresponding AO–WF coefficient and index $\boldsymbol{\mu}$ combines the indices μ and $\mathbf{R}$. We define as the variational functional to be maximized, the Löwdin population of the given localized function on a set of atoms surrounding its centroid:

$$I_i^{(\mathrm{no})} = \sum_{\boldsymbol{\mu}\in\mathcal{S}_i^{(\mathrm{loc})}} (C_{\boldsymbol{\mu},i}^{(\mathrm{no}-Lw)})^2 = \langle W_i^{(\mathrm{no})} | \hat{P}_{\mathcal{S}_i^{(\mathrm{loc})}}^{(Lw)} | W_i^{(\mathrm{no})} \rangle, \tag{8}$$

where $C_{\mu,i}^{(\text{no}-Lw)}$ are the expansion coefficients of the NOLO $W_i^{(\text{no})}(\mathbf{r})$ over the Löwdin-orthogonalized atomic basis set functions $\phi_\mu^{(Lw)}(\mathbf{r})$:

$$W_i^{(\text{no})}(\mathbf{r}) = \sum_\mu C_{\mu,i}^{(\text{no}-\text{Lw})} \phi_\mu^{(Lw)}(\mathbf{r}), \tag{9}$$

$$\phi_\mu^{(Lw)}(\mathbf{r}) = \sum_{\mu'} (S_{\mu'\mu}^{(\text{AO})})^{-1/2} \phi_{\mu'}(\mathbf{r}), \tag{10}$$

$S_{\mu,\mu'}^{(\text{AO})}$ is the AO overlap matrix, and $\hat{P}_{\mathcal{S}_i^{(\text{loc})}}^{(Lw)}$ is an operator projecting onto the space of the Löwdin-orthogonalized AOs from the given set of atoms:

$$\hat{P}_{\mathcal{S}_i^{(\text{loc})}}^{(Lw)} = \sum_{\mu \in \mathcal{S}_i^{(\text{loc})}} |\phi_\mu^{(Lw)}\rangle\langle\phi_\mu^{(Lw)}|. \tag{11}$$

The set of atomic orbitals $\mathcal{S}_i^{(\text{loc})}$ determines the region (or the set of atoms), the corresponding localized functions $W_i^{(\text{no})}(\mathbf{r})$ should concentrate within. The functional Eq. (8) acts in a way similar to the one proposed in Reboredo and Williamson [23] and, at the same time, is appropriate for LCAO calculations.

Maximization of the functional (8) with a constraint of normalization of the functions $W_i^{(\text{no})}(\mathbf{r})$ can be carried out by diagonalizing the matrix of the operator $\hat{P}_{\mathcal{S}_i^{(\text{loc})}}^{(Lw)}$ (11). The eigenvector of this matrix, corresponding to the highest eigenvalue, holds the coefficients of the expansion (9), providing the most localized NOLO in the sense of the functional $I_i^{(\text{no})}$. The basis for this matrix could be formed by the Bloch functions of the occupied bands [23,29], WFs or even atomic orbitals, projected onto the occupied space [17].

During generation of the nonorthogonal functions one should not lose the rank of the space spanned by these functions (i.e., the linear independency of the functions should be kept). For the crystals of sufficiently high symmetry the rank can be preserved automatically, if the basis used for the variational procedure is symmetrized [8]. The localized functions are basis functions of irreducible representations (irreps) of the site groups related to the centroids of these functions [7]. Thus, if the basis for the functional (8) is symmetrized according to these very site group and irrep, the centering site of the localized function cannot change [7,8]. So, when the localized functions are centered on an isolated symmetry point, the symmetrization of the basis keeps the functions linearly independent. Moreover, even if the centroid of the localized function is not an isolated symmetry point, but the symmetry of the crystal is still high enough, the rank still may be kept by just taking symmetry into account. Yet, in this case the latter is not guaranteed. If preservation of rank cannot be assured by applying symmetry, one may have to apply some additional constraints — e.g., as suggested in Lui et al. [15], the constraint of the nonsingularity for the localized functions overlap matrix, or Feng et al. [10], the constraint of freezing the centroids of the localized functions. However, the latter constraint might collide with the requirement of higher localization, if the centers of the functions are not fixed by the symmetry.

The variational procedure should be performed in the space of the first vectors of the corresponding irreps. In other words, only symmetry-unique localized functions are to be obtained variationally. The remaining functions can be generated by applying the appropriate symmetry operators [8, 7,29]. If there is more than one independent set of localized functions centered on the same site and transforming according to the same irrep, one can obtain all of them in the variational procedure either by taking the eigenfunctions of the diagonalized matrix corresponding to several of the highest eigenvalues instead of just one, or by performing independent procedures for their generation. In the latter case, the variational procedure should be run in the WF-basis, and only one out of several WFs with the considered symmetry, which are centered on the same site, should be included into this basis. The symmetry properties of the WFs and localized functions can either be determined before (or without) the actual construction of WFs, by analyzing the symmetry of the Bloch functions corresponding to the valence bands [7,29], or a posteriori by applying the technique of Casassa et al. [3] to nonsymmetrized WFs [34].

3 Fitting of localized functions

Unlike molecular localized orbitals, WFs have an infinite support, which means that the actual limit for the summation over the index $\mathbf{R}$ in (7) is infinity. Thus, in numerical calculations the sum (7) has to be truncated according to some threshold. In the present LMP2 code implemented in the `CRYSCOR` program, [22] a screening based on the magnitude of the WF coefficients is employed. In effect, all contributions due to WF coefficients below that screening threshold are disregarded. Unfortunately, only relatively loose thresholds can presently be afforded in practical calculations [22] and the numbers of $\mathbf{R}$ taken into account may become a critical parameter influencing significantly the accuracy of the calculation. For example, the four-index transformation of the two-electron integrals in `CRYSCOR` uses tolerances for the LCAO coefficients of the WFs of 10^{-2}–10^{-3}. Such loose thresholds can hardly guarantee sufficient accuracy for reliable results, yet tighter thresholds render the calculations too expensive.

In order to treat the truncation of the WFs' tails in an improved way we introduce the criterion $I_i^{(\text{fit})}$:

$$I_i^{(\text{fit})} = \int (\widetilde{W}_i(\mathbf{r}) - W_i(\mathbf{r}))^2 d\mathbf{r}. \tag{12}$$

Here the function $\widetilde{W}_i(\mathbf{r})$ is the truncated localized function:

$$\widetilde{W}_i(\mathbf{r}) = \sum_{\mu \in \mathcal{S}_i^{(\text{fit})}} C_{\mu,i} \phi_\mu(\mathbf{r}), \tag{13}$$

and $\mathcal{S}_i^{(\text{fit})}$ contains an incomplete set of atomic functions centered only on atoms, which are in a way close to the centroid of the WF $W_i(\mathbf{r})$. The criterion (12) is utilized to *fit* the

WF spanned by the basis functions within $\mathcal{S}_i^{(\mathrm{fit})}$ (WF–AO fit domain) in a least-square sense with respect to the untruncated function. Hence, minimizing the functional $I_i^{(\mathrm{fit})}$ yields the fitted WFs

$$\widetilde{W}_i^{(\mathrm{fit})}(\mathbf{r}) = \sum_{\mu \in \mathcal{S}_i^{(\mathrm{fit})}} C_{\mu,i}^{(\mathrm{fit})} \phi_\mu(\mathbf{r}), \tag{14}$$

with the fitted coefficients $C_{\mu,i}^{(\mathrm{fit})}$, obtained by solving the linear equations

$$\sum_{\mu' \in \mathcal{S}_i^{(\mathrm{fit})}} S_{\mu,\mu'}^{(\mathrm{AO})} C_{\mu',i}^{(\mathrm{fit})} = \sum_{\mu'} S_{\mu,\mu'}^{(\mathrm{AO})} C_{\mu',i}. \tag{15}$$

The fitting according to (15) is formally equivalent to other fitting techniques like density fitting [26,32], with the AOs from the WF–AO fit-domain in the role of the auxiliary basis functions.

By successively adding coordination spheres of atoms and thus their atomic functions to the set $\mathcal{S}_i^{(\mathrm{fit})}$ and refitting the coefficients $C_{\mu',i}^{(\mathrm{fit})}$ one can finally obtain the fitted WF $\widetilde{W}_i^{(\mathrm{fit})}(\mathbf{r})$ which to a given tolerance approximates the exact WF $W_i(\mathbf{r})$. Within a chosen tolerance the number of atoms contributing significantly to the fitted WFs $\widetilde{W}_i^{(\mathrm{fit})}(\mathbf{r})$ is essentially smaller than that of the functions $\widetilde{W}_i(\mathbf{r})$.

The fitting procedure can be applied also to the NOLOs, once they are constructed. The localization-fitting process for the NOLOs is most efficient when the set of atomic orbitals $\mathcal{S}_i^{(\mathrm{loc})}$ and WF–AO fit-domain $\mathcal{S}_i^{(\mathrm{fit})}$ at both stages are chosen to be the same. Indeed, small values of the fitting functional $I_i^{(\mathrm{fit})}$ can be expected, if all the sites with relatively large value of the NOLO are included in the corresponding WF–AO fit-domain. Thus, it does not make sense to use domains $\mathcal{S}_i^{(\mathrm{fit})}$ smaller than $\mathcal{S}_i^{(\mathrm{loc})}$, since the latter guarantees small values for localized functions only outside its region. On the other hand, if the value for the fitting functional within the chosen WF–AO fit-domain $\mathcal{S}_i^{(\mathrm{fit})}$ is not sufficiently small, it is more efficient, when one does not just enlarge it (as in the case of WFs), but recalculates beforehand the reference NOLOs with the new set $\mathcal{S}_i^{(\mathrm{loc})}$ enlarged accordingly.

Actually, only the LCAO coefficients of the NOLO (or WFs) $C_{\mu,i}^{(\mathrm{no})}$ ($C_{\mu,i}$) can be refitted by minimizing the functional (12). The Löwdin coefficients $C_{\mu,i}^{(\mathrm{no}-Lw)}$ ($C_{\mu,i}^{(Lw)}$) cannot be modified by the fitting (15), since the corresponding AO overlap matrix involved is the identity matrix. It means that the Löwdin coefficients within the WF–AO fit-domain are optimal with respect to the functional (12) providing its minimum. For orthogonal atomic-like basis sets (such as Löwdin-orthogonalized AOs, or orthogonal wavelets [4]) the fitting process just corresponds to the truncation of the WF coefficients according to the chosen WF–AO fit-domain. Constructing the NOLO in the case of orthogonal basis sets might be more efficient, since the localization functional (8) addresses the coefficients relative to the orthogonalized basis set functions directly. The efficiency of the fitting procedure for the localized functions depends on redundancies carried in the corresponding basis sets, which are large in case of highly overlapping basis sets.

Summarizing the aforesaid, the method for obtaining ultralocalized nonorthogonal fitted functions implies the following. Once the symmetry of the localized orbitals has been determined, for every symmetry-unique function the variational procedure in the symmetrized basis according to the functional (8) is performed, followed by generating the fitted functions with limited support from the former by solving (15). The sets of the atomic orbitals $\mathcal{S}_i^{(\mathrm{loc})}$ and $\mathcal{S}_i^{(\mathrm{fit})}$ in these two steps are to be taken the same. If the value of the fitting functional (12) doesn't drop below a predefined threshold, a new star of atoms is added and the process is repeated.

The atomic functions projected onto the orthogonal complement of the occupied space (PAOs), which represent in the framework of local correlation methods the virtual states [22,25], are also localized functions. Therefore, similar techniques could also be used to restrict the spatial extent of these PAOs. However, since PAOs are to be constructed by projecting out the space spanned by the *fitted* WFs or NOLOs (rather than the untruncated parental WFs or NOLOs), we anticipate that PAOs will automatically posses a limited support, determined basically by the WF AO fit-domains $\mathcal{S}_i^{(\mathrm{fit})}$.

4 Calculations

The proposed methods have been applied to the diamond crystal. This crystal has a face-centered cubic (fcc) lattice with the space group O_h^7. The four valence WFs per unit cell are centered in the middle of the C–C covalent bonds. These sites are isolated symmetry points with D_{3d} site-symmetry group. The WFs and nonorthogonal functions studied below are linear combinations of HF canonical crystalline orbitals, obtained by the periodic LCAO code `CRYSTAL03` [6]. The optimized 6-21G* basis set [5] was used for the calculations. The reciprocal space was sampled by 8×8×8 Monkhorst–Pack [18] k-points mesh.

Figure 1 gives the orthogonal (dashed line) and nonorthogonal localized orbitals, obtained with the sets $\mathcal{S}^{(\mathrm{loc})}$ of three (dash-dotted line) and five (solid line) stars of atoms. It is clearly seen that, when the chosen $\mathcal{S}^{(\mathrm{loc})}$ is sufficiently large (three and, especially, five stars), the nonorthogonal functions are essentially more localized and smooth than the orthogonal ones. They do not posses the oscillatory tails, as their values drop virtually to zero beyond a certain region.

Though the explicit behavior of the NOLOs is quite clear, their higher concentration around the centering point might not be noticed through analyzing their LCAO coefficients. The values of the Löwdin coefficients $C_{\mu,i}^{(Lw)}$ and $C_{\mu,i}^{(\mathrm{no}-Lw)}$ of orthogonal (circles) and nonorthogonal (stars) localized functions, averaged over AOs of different stars of atoms are given in Fig. 2a. As one might expect, the coefficients for the nonorthogonal functions on average are considerably smaller (one or more orders of magnitude) than the ones of the WFs. But the LCAO coefficients, $C_{\mu,i}$ and $C_{\mu,i}^{(\mathrm{no})}$, which are to be used in the actual calculations, behave differently. Their averaged

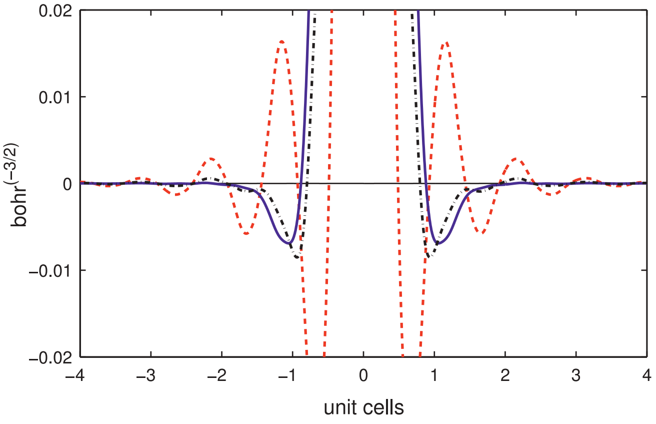

Fig. 1 Wannier function(*dashed*), nonorthogonal localized orbitals localized within three (*dash-dotted*) and five (*solid*) stars of atoms, for diamond along the [001] axis. The origin is placed at the WF's centroid

values are presented in Fig. 2b. As seen from this plot, the coefficients for WFs and NOLOs are approximately of the same order of magnitude. This fact indicates that the absolute values of the LCAO coefficients cannot serve as an efficient criterion for dropping the tails of the localized functions. The coefficients may have relatively high absolute values and at the same time the corresponding AOs might cancel each other due to redundancies in the nonorthogonal AO basis. This annoying effect is properly cured by introducing the criterion (12) and subsequent fitting of the localized functions. This criterion reflects more the explicit WF's or NOLO's behavior rather than the behavior of its coefficients. As a result of the fitting, the coefficients of otherwise canceling AOs would be set to zero, while the overall precision is maintained.

The values for the functional $I_i^{(\text{fit})}$ (12) for functions of four types – non-fitted and fitted WFs and non-fitted and fitted NOLOs – are given in Fig. 3 on logarithmic scale. It is evident that fitted WFs, and especially fitted NOLOs, when used in the calculations of the two-electron integrals, can significantly improve the performance. Thus, for a chosen threshold of the functional (12) (which, at the end of the day, reflects the accuracy in the correlation energy), switching to the fitted and/or nonorthogonal functions significantly curtails the number of atoms over which these function extend, which in turn implies a substantial reduction in the computational cost of local correlation calculations employing these functions. For example, the WF fitted to the WF–AO fit-domain comprising the AOs from 8 atoms has approximately the same accuracy (the value of the functional (12)) as the non-fitted WF with the support of the AOs from not less than 32 atoms. Similarly, the value of the fitting functional for a NOLO using a WF–AO fit-domain of 32 atoms has approximately the same value as that of a WF with a support of 136 atoms, etc. Alternatively, if the localization region (and with this the cost of the correlation calculation) is fixed, the accuracy may be essentially increased by up to two orders of magnitude (at least at the level of the localized functions). These two strategies can also be combined.

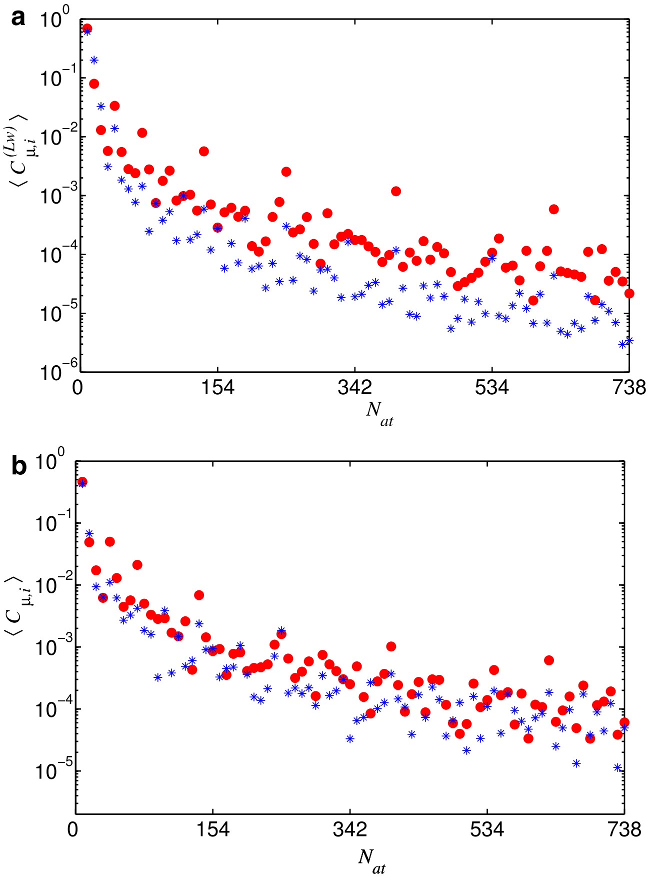

Fig. 2 Averaged over the atomic orbitals of the given star of atoms the Löwdin (**a**) and LCAO (**b**) coefficients of the WF (*circles*) and NOLO (*stars*) for diamond. The NOLO was localized within five stars of atoms. The axis of abscissas is partitioned according to the indices of the stars, but labeled by the number of atoms within these stars

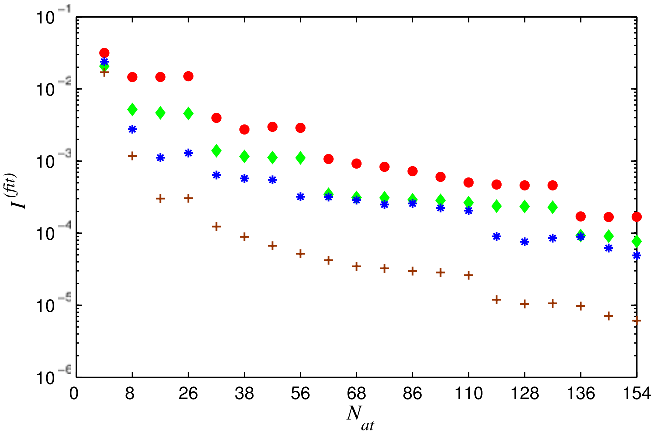

Fig. 3 The values of the functional (12) for the WF (*circles*), fitted WFs (*rhombs*), NOLOs (*stars*), fitted NOLOs for diamond. The axis of abscissas is partitioned according to the indices of stars included in the calculation of this functional (and also used for fitting and construction of the corresponding WFs and NOLOs), but labeled by the number of atoms within these stars

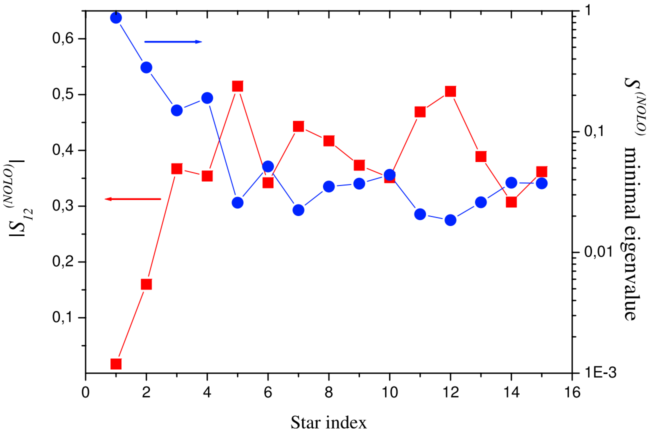

Fig. 4 The minimal eigenvalue (*circles*) of the NOLO's overlap matrix, and the absolute value of the actual overlap between the closest NOLOs (*squares*) in diamond

Table 1 Comparison of the MP2 calculations for diamond using nonfitted Wannier functions (WFs) truncated beyond three stars of atoms and fitted WFs with the WF–AO fit-domain, consisting of three stars

Type of calculation	MP2 energy E_2 (Hartrees)	$E_2 - E_2^{(\mathrm{ref})}$ (Hartrees)	CPU time (sec)
Reference	−0.22362		198626
Nonfitted WF	−0.21652	0.00710	21138
Fitted WF	−0.22238	0.00124	21272

The reference calculation $E_2^{(\mathrm{ref})}$ corresponds to the lowest manageable by the `CRYSCOR` program (in its current status) threshold (0.0005) for the WFs coefficients. The calculations are rather illustrative, since all other thresholds were taken to be quite large in order to speed up the calculation

The axis of abscissas in Figs. 2 and 3 is partitioned equidistantly with respect to different stars of atoms. This implies that the distance between the atoms of a star and the center does not increase linearly with the star index. As a consequence the WF's coefficients and the functional (12) values tend towards a plateau (in the logarithmic scale) instead of decreasing linearly to zero. It can also be seen from Fig. 3 that the non-fitted NOLOs for some sets $\mathcal{S}^{(\mathrm{loc})}$ correspond to approximately the same values of the integral $I_i^{(\mathrm{fit})}$ as the fitted WFs.

Figure 4 contains important information about the overlap and linear independence of the nonorthogonal functions. Evidently, while the set $\mathcal{S}^{(\mathrm{loc})}$ is relatively small (one star), the corresponding nonorthogonal localized functions do not overlap essentially and also do not differ much from the corresponding orthogonal WFs. At larger sets $\mathcal{S}^{(\mathrm{loc})}$ the nonorthogonal functions become noticeably different – their localization and overlap essentially increase and the minimal eigenvalue of the NO–LOs overlap matrix ($S^{(\mathrm{NOLO})}$) becomes smaller. The maximal compactness of the functions is achieved for the sets $\mathcal{S}^{(\mathrm{loc})}$ comprising five stars. The minimal eigenvalue of the overlap matrix then drops to a value somewhat higher than 0.01, so these functions remain linearly independent. Further increase of the set $\mathcal{S}^{(\mathrm{loc})}$ does not lead to any major changes in the studied parameters.

As mentioned above, fitting of the WFs and NOLOs according to (15) can improve the criterion (12) only if the basis functions are overlapping. Calculations with the STO-3G minimal basis set revealed the Löwdin-like behavior of the WFs and NOLO coefficients, as the LCAO coefficients explicitly reflected the improved localization of the nonorthogonal functions. Moreover, fitting of the localized function coefficients did not substantially reduce the functional's value, just as in the case of Löwdin coefficients before. The reason of this behavior is that the STO-3G basis does not contain diffuse functions and carries less redundancy than the 6-21G* basis set. The overlap of the AOs is small, hence, the Löwdin orthogonalization has relatively little effect on the AOs. We can thus conclude that the more diffuse orbitals are included in the basis set, the greater is the impact of the fitting on the localized functions' coefficients.

Finally, the proposed techniques were examined in illustrative LMP2 calculations. Since in the present version of the `CRYSCOR` code a possible nonorthogonality of the localized functions, representing the occupied manifold is not taken into account, the LMP2 energy in diamond has been calculated using only orthogonal WFs with and without fitting. The weak and distant pair distances were set to 2 Å [22]. The threshold for PAOs coefficients was set to 10^{-2}. Only the coefficients for the AOs from the first three stars of atoms were included in the transformation of the 4-index integrals (WF fitting was also performed within the first three stars of atoms). LMP2 correlation energy and timings were then compared with the corresponding values of a reference calculation, which were obtained by employing quite a low screening threshold of $5 \cdot 10^{-4}$ for the WF coefficients (smaller thresholds are hardly manageable, presently [22]). The results of these calculations are presented in Table 1. Evidently, the fitting of WF coefficients can improve the accuracy and/or reduce the computational cost of the LMP2 calculations for solids. As seen from Table 1, if fitted WFs are used, the estimate of the error in the calculated MP2 energy (the difference between the calculated and reference energies) becomes substantially smaller, while the CPU time of the calculation remains about the same.

5 Conclusions

In conclusion, a method for generating compact nonorthogonal ultralocalized functions has been devised. The proposed procedure is to maximize the Löwdin population of the functions within a given set of atomic functions. It turns out that this set of AOs has to reach a certain size, otherwise no significant difference relative to the parental orthogonal WFs, i.e., not an increased compactness is achieved. Furthermore, a scheme for fitting truncated WFs and nonorthogonal localized orbitals (support of AOs limited to WF–AO fit-domains) to the original untruncated functions has been proposed. The calculations showed that by virtue of the fitting process the truncated fitted localized functions, and in particular the non-orthogonal ones, improve insofar that the specified criterion considerably decreases compared to the corresponding functions obtained by just truncating accordingly. Preliminary calculations of the LMP2 energy using the LMP2 method implemented in the `CRYSCOR` program demonstrated that an improvement in the accuracy of the calculation can be achieved if fitted rather than just truncated WFs are employed.

Acknowledgements The authors acknowledge the support by DFG-grant SCHU 1456/3-1.

References

1. Anderson PW (1968) Phys Rev Lett 21:13
2. Calzolari A, Marzari N, Souza I, Nardelli MB (2004) Phys Rev B 69:035108
3. Casassa S, Zicovich-Wilson CM, Pisani C (2005) Theor Chem Acc (to be published)
4. Clow SD, Johnson BR (2003) Phys Rev B 68:235107
5. `CRYSTAL` basis set database http://www.crystal.unito.it/basis.html
6. Dovesi R, Saunders VR, Roetti C, Causa M, Harrison NM, Orlando R, Apra E (2003) `CRYSTAL` 2003 User's Manual. Torino University, Torino
7. Evarestov RA, Smirnov VP (1997) Site symmetry in crystals: theory and applications. Springer series in solid state sciences, Vol 108. Springer, Berlin Heidelberg New York
8. Evarestov RA, Usvyat DE, Smirnov VP (2003) Phys Solid State 45:2072
9. Evarestov RA, Smirnov VP, Usvyat DE (2003) Solid State Commun 127:423
10. Feng H, Bian J, Li L, Yang W (2004) J Chem Phys 120:9458
11. Goedecker S (1999) Rev Mod Phys 71:1085
12. He L, Vanderbilt D (2001) Phys Rev Lett 86:5341
13. King-Smith RD, Vanderbilt D (1993) Phys Rev B 47:1651
14. Kohn W (1959) Phys Rev 115:809
15. Lui S, Perez-Jorda JM, Yang W (2000) J Chem Phys 112:1634
16. Marzari N, Vanderbilt D (1097) Phys Rev B 56:12847
17. Mayer I, Räther G, Suhai S (1998) Chem Phys Lett 293:81
18. Monkhorst HJ, Pack JD (2003) Phys Rev B 13:5188
19. Noel Y, Zicovich-Wilson CM, Civallery B, D'Arco Ph, Dovesi R (2001) Phys Rev B 65:014111
20. Paulus B, Rosciszewski K, Stoll H, Birkenheuer U (2003) Phys Chem Chem Phys 5(24):5523
21. Paulus B (2003) Chem Phys Lett 371:7
22. Pisani C, Busso M, Cappecchi G, Casassa S, Dovesi R, Maschio L, Zicovich-Wilson C, Schütz M (2005) J Chem Phys 122:094113
23. Reboredo FA, Williamson AJ (2005) Phys Rev B 71:121105(R)
24. Sagui C, Pomorski P, Darden TA, Roland C (2004) J Chem Phys 120:4530
25. Schütz M, Hetzer G, Werner H-J (1999) J Chem Phys 111:5691
26. Schütz M, Manby FR (2003) Phys Chem Chem Phys 5:3349
27. Shukla A, Dolg M, Fulde P, Stoll H (1999) Phys Rev B 60:5211
28. Skylaris C-K, Mostofi AA, Haynes PD, Diguez O, Payne MC (2002) Phys Rev B 66:035119
29. Smirnov VP, Usvyat DE (2001) Phys Rev B 64:245108
30. Sporkmann B, Bross H (1994) Phys Rev B 49:10869
31. Wannier G (1937) Phys Rev 52:191
32. Werner H-J, Manby FR, Knowles PJ (2003) J Chem Phys 118:8149
33. Zeiner P, Dirl R, Davies BL (1998) Phys Rev B 58:7681
34. Zicovich-Wilson CM, Dovesi R, Saunders VR (2001) J Chem Phys 115:9708
35. Zicovich-Wilson CM, Bert A, Roetti C, Dovesi R, Saunders VR (2002) J Chem Phys 116:1120