Frequency localization properties of the density matrix and its resulting hypersparsity in a wavelet representation

S. Goedecker

Max-Planck Institute for Solid State Research, Stuttgart, Germany

O. V. Ivanov

P. N. Lebedev Physical Institute, Moscow, Russia

(Received 18 September 1998)

$O(N)$ methods are based on the localization properties of the density matrix in real space, an effect refered to as nearsightedness. We show that, in addition to this real-space localization there is also a localization in Fourier space. Using a basis set with good localization properties in both real and Fourier space such as wavelets, one can exploit both localization properties to obtain a density matrix that exhibits additional sparseness properties compared to the scenario where one has a basis set with real-space localization only. Taking advantage of this hypersparsity, it is possible to represent very large quantum-mechanical systems in a highly compact way. This can be done both for insulating and metallic systems. We expect that hypersparsity will pave the way for highly accurate $O(N)$ calculations of large systems requiring many basis functions per atom. [S0163-1829(99)02904-5]

Methods for the calculation of the electronic structure that exhibit linear scaling with respect to the size of the system, so-called $O(N)$ methods,[1] are an important topic in physics and chemistry. Only with this kind of algorithms it is possible to calculate the properties of very large systems that are relevant in many practical applications. $O(N)$ methods have become practically a standard for large-scale tight-binding calculations where the ratio of the number of basis functions to the number of electrons is of the order of 2. In highly accurate density-functional-type calculations, where the number of basis functions is very large compared to the number of electrons, these methods have on the other hand not been widely used up to now. This comes from the fact that $O(N)$ methods are only efficient if the density matrix is very sparse. Large conventional basis sets lead, however, to density matrices that are not sufficiently sparse. The sparseness of the density matrix in connection with conventional localized basis sets is due to the real-space localization properties of the density matrix, the so-called real-space nearsightedness[2] of the system. We will show that the density matrix has in addition also Fourier space localization properties and that this fact can be used to make the density matrix significantly more sparse. These gains will be particularly important if the density is not strongly localized in real space, which is the case for small-band-gap and metallic systems.

For reasons of simplicity and in order to be able to do certain numerical calculations without any truncation we will concentrate in the following on the one-dimensional case even though all the principles are applicable to the three-dimensional case. For our investigations we use a test system that is characterized by a simple harmonic potential $-2\sin(2\pi r)$ and we applied periodic boundary conditions. The valence band extends from roughly 0 to 4 atomic units and it is followed by a rather large gap of roughly two atomic units. By occupying each primitive cell that has a length of one atomic unit with one electron pair one thus obtains an insulator, while one obtains a metal by assigning, for instance, two electron pairs to each cell.

The density matrix F in a real space representation can be modeled by

$$F(r,r') = C\exp(\kappa r)\frac{\sin\left[\frac{\pi}{a}(r-r')\right]}{r-r'}. \tag{1}$$

This is in principle only the asymptotic form, but is also a good approximation for small values of r and r'. The decay constant κ grows with increasing gap and is zero in a metallic system.[3] The oscillation length a is of the order of the average distance among the valence electrons in a metal and of the order of the interatomic spacing in an insulator. C is a normalization constant. Evidently $F(r,r')$ has both a typical length scale of $1/\kappa$ and a dominating Fourier component of π/a. The localization in real space will be particularly good for strongly insulating systems whereas the localization around the typical Fourier component becomes better for decreasing gap sizes. It is well known that in most realistic materials $F(r,r')$ decays over several oscillation lengths indicating a combined localization in real and Fourier space.

Let us now briefly review some essential facts from wavelet theory.[4] For simplicity we will work with the standard Daubechies wavelet family.[4] There are two sets of fundamental functions, the scaling functions ϕ and the wavelets ψ. The scaling functions are essentially ordinary localized functions comparable, for instance, to Gaussians in many respects. Their advantage over Gaussians is, however, that they are orthogonal. To form a basis set one has to take all the integer translations i of the scaling functions at a certain resolution level k,
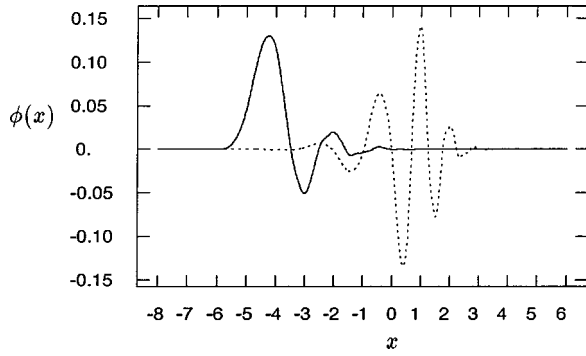
$$\phi_i^k(x) = \phi(2^k x - i). \tag{2}$$

FIG. 1. The Daubechies scaling function (solid line) and wavelet (dashed line) used in this work.

Applying the unitary fast wavelet transform, one can transform the set of the scaling functions ϕ into the equivalent set of wavelets ψ. The wavelet basis one obtains in this way consists now of translated wavelets at different resolution levels. What is important in this context is that the scaling functions do not exhibit any localization in Fourier space whereas the wavelets do exhibit this feature. The Fourier spectrum for the Daubechies wavelet of degree 12 (Fig. 1) that was used in this work is shown in Fig. 2.

Let us now go over to the $O(N)$ context. The most general $O(N)$ methods are the ones that calculate the full density matrix. In these methods one calculates the matrix elements of F with respect to a basis set χ_i:

$$F_{i,j}= \int \int \chi_i(r)F(r,r')\chi_j(r')drdr'. \quad (3)$$

Let us first discuss the case where the χ's are ordinary localized functions. In order to be able to do a fair evaluation of the advantages obtained by additional Fourier localization we will choose as our set of localized functions the set of scaling functions that is completely equivalent to the set of wavelets used in the comparison. It is obvious that the width of the scaling functions has to be less than the wavelength of the oscillatory part of F in order to be an accurate basis. This means that

$$F_{i,j}= \int \int \phi_i^k(r)F(r,r')\phi_j^k(r')drdr' \approx F(R_i,R_j), \quad (4)$$

where R_i is the position of the ith scaling function. The function $F(0,r)$ for our test system is shown in Fig. 3.
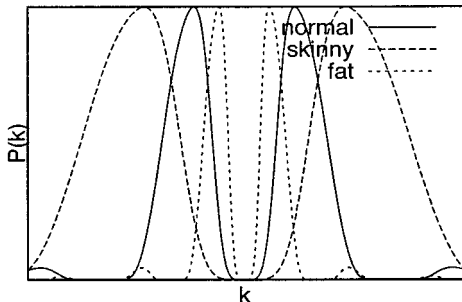

FIG. 2. The Fourier power spectrum $P(k)$ of the Daubechies wavelet at three different resolution levels denoted by skinny, normal and fat. As one sees the peaks at the dominating Fourier components are reasonably well separated from each other.
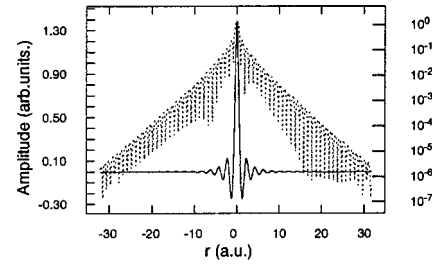

FIG. 3. The function $F(0,r)$. Along the left y-axis going with the solid line it is plotted on a normal scale and along the right y-axis going with the dashed line on a logarithmic scale.

Inspite of the exponential decay, one has to allow for a fairly large subvolume of the whole 64-atom system before one can truncate $F(0,r)$ without a large truncation error. This means, that most of the elements of the density matrix have to be calculated and that therefore any traditional $O(N)$ scheme will not be very efficient. The structure of the density matrix using four scaling functions per atom is shown in Fig. 4.

Let us now look at the same matrix in a wavelet basis, where we have to calculate matrix elements of the type

$$\int \int \psi_i^k(r)F(r,r')\psi_j^{k'}(r')drdr'. \quad (5)$$

Note that in contrast to Eq. (4) there are now matrix elements between wavelets at different resolution levels k and k', that have different dominating Fourier components. In accordance with the usual convention we order the wavelets at the different resolution levels such that the skinny wavelets follow the fat ones. So, in our example, the indices 129 to 256 refer to the thinnest wavelets, the indices 65 to 128 to the not quite so thin wavelets, etc. Let us first look at the upper right 128 times 128 block in Fig. 5. This block represents only matrix elements among the most thin wavelets. From Fig. 4 we would expect to find elements larger than 1.e-2 and that the bandwidth is close to half of the size of the block. In reality all elements are smaller than 1.e-2, and the bandwidth is much less than the one in Fig. 4. This is the effect of the above postulated hypersparsity. If both basis functions in Eq. (5) are thin wavelets, the function $F(r,r')$ cannot couple them since its dominating wavelength is much larger than the dominating wavelength of the two wavelets. As a matter of fact, any matrix element will be strongly damped unless all the three terms in Eq. (5) have comparable wavelengths. Therefore, also, all the other fingers in Fig. 4 that represent coupling between wavelets of different stature (i.e., different k and k') have a small bandwidth. Because their spatial frequencies do not match, the decay for wavelets positioned a certain distance apart is much faster than one would expect from the decay behavior of the amplitude of $F(r,r')$ alone.

To quantitatively assess the advantages of the additional Fourier localization, compared with spatial localization only, let us plot the number of coefficients that are necessary to represent the density matrix with a certain error. We define the error as the 2-norm of the difference between the exact and the truncated density matrix. As one can see from Fig. 6 the effect is dramatic. The accuracy obtained with the same number of coefficients with the wavelet representation is sev-
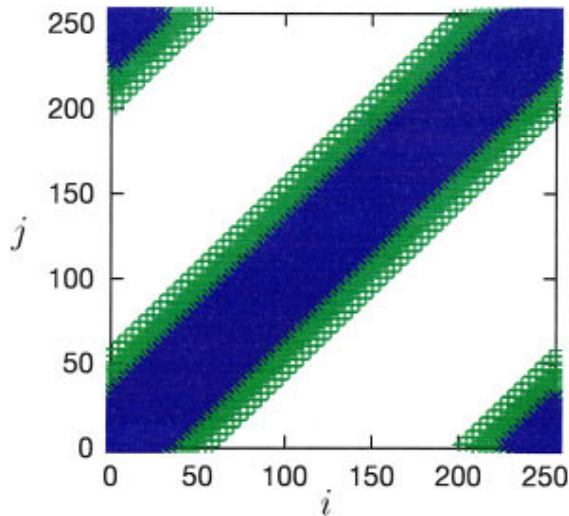
FIG. 4. (Color) Structure of the density matrix in a scaling function representation. Elements larger than 1.e-3 are denoted by green areas, elements larger than 1.e-2 by blue areas.

eral orders of magnitude smaller than the one obtained with the scaling function representation.

In the case of an insulator at zero electronic temperature, the density matrix can compactly be represented in terms of Wannier functions. The Wannier functions are characterized by a spatial decay length and a dominant Fourier component as well. So the same principles apply and a wavelet representation is significantly more efficient than a scaling function representation (Fig. 6). Since one needs only one Wannier function per electron pair instead of all the columns of the density matrix the prefactor in the Wannier representation is, however, smaller.

Let us next consider the metallic systems. For the case of jellium, plane waves would obviously give the most compact representation, leading to a strictly diagonal density matrix. So the principle of spatial nearsightedness does not apply and the Fourier space localization of the basis functions will
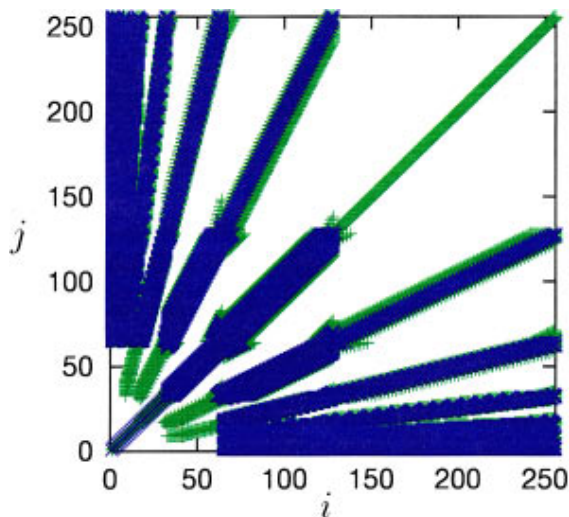

FIG. 5. (Color) Structure of the density matrix in a wavelet representation. The convention for the colors is the same as in Fig. 4. The decay is so fast in this case that the green area around the blue area is so thin to be hardly visible in some regions.
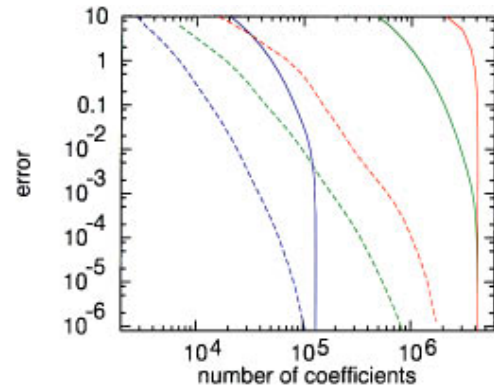

FIG. 6. (Color) The error in the density matrix versus the size of the data set necessary for its representation. Solid lines correspond to a scaling function representation, dashed lines to a wavelet representation. The two red curves are for a metallic system, the two green curves for an insulator and the two blue curves for an insulator where the density matrix was constructed indirectly via the Wannier functions.

even be more important than the real space localization in the metallic case. As one sees from Fig. 1 our wavelets have a more pronounced spatial localization than Fourier localization. Other families of wavelet that have a more pronounced Fourier space localization might in this context be more efficient than the Daubechies wavelet used in these tests. The structure of the matrix is shown in Fig. 7 for the case of the wavelet representation of the test system containing 128 electron pairs and 512 basis functions. In the case of a scaling function representation all elements are larger than 1.e-2 and the corresponding matrix is therefore not shown.

As expected, the quantitative evaluation in Fig. 6 shows that the savings of a wavelet representations compared to a scaling function representation are even more important in the metallic case than in the case of an insulator. In both representations more coefficients are, however, needed than in the insulating case. This comes from the fact that the decay behavior of matrix elements with moderate spatial frequency disparity is still determined by the slow spatial decay of $F(r,r')$. In addition to the finger structures there are therefore, some full blocks in Fig. 7.
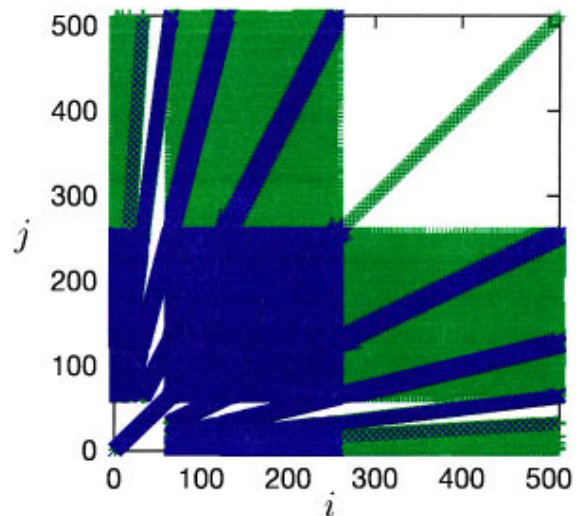

FIG. 7. (Color) Same as Fig. 5 but for a metallic system.

In the context of large systems, the scaling of the required number of basis functions with respect to system size is of course important. As is well known, even with basis sets that have real-space localization only, this scaling is linear. Hence it is to be expected to be linear as well in the case of a combined spatial and Fourier localization. This assumption was clearly confirmed by our numerical experiments where we studied model systems containing between 16 and 128 ''atoms.'' For metallic systems ordinary localized basis functions give a quadratic scaling unless one goes over to unrealistically large systems. Our numerical test for metallic systems containing between 16 and 128 ''atoms'' gave an exponent of 1.66. So the advantages of basis sets with frequency localization will grow with increasing system sizes.

Like in the majority of wavelet based calculations,[5,6] the basic computational kernels of the calculations presented here were implemented with respect to a scaling function basis. The final results were then transformed into a wavelet basis to examine the influence of truncation. In order to take full advantage of a truncated wavelet basis set, it would be necessary to do the whole calculation within the truncated wavelet basis. Matrix times vector multiplications are the basic step of several $O(N)$ schemes and it has been demonstrated how they can be implemented with strictly linear scaling with respect to the number of nonzero coefficients.[7] So the basic observation of this paper could directly be used to reduce the numerical effort in the same way as the number of necessary basis functions.

Based on the observation that the density matrix $F(r,r')$ is characterized by both a typical length scale and a dominating Fourier component, we have demonstrated that quantum mechanical systems exhibit not only a near sightedness in real space but also a near sightedness in Fourier space. A basis set that has combined localization in both real and Fourier space can therefore significantly reduce the number of data that are needed to describe it. We expect that this phenomenon that we call hyper-sparsity will allow us to overcome the present limitation of $O(N)$ scheme to systems where the number of basis functions per atom is rather small.

---

[1] S. Goedecker, Rev. Mod. Phys. (to be published July 1999); W. Yang, Phys. Rev. Lett. **66**, 1438 (1991); W. Yang and T-S. Lee, J. Chem. Phys. **103**, 5674 (1995); S. Goedecker and M. Teter, Phys. Rev. B **51**, 9455 (1995); A. F. Voter, J. D. Kress, and R. N. Silver, *ibid.* **53**, 12 733 (1996); S. Goedecker, J. Comput. Phys. **118**, 261 (1995); X.-P. Li, W. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10 891 (1993); M. S. Daw, *ibid.* **47**, 10 895 (1993); **6**, 9153 (1994); F. Mauri, G. Galli, and R. Car, *ibid.* **47**, 9973 (1993); P. Ordejon, D. A. Drabold, M. P. Grumbach, and R. M. Martin, *ibid.* **48**, 14 646 (1993); J. Kim, F. Mauri, and G. Galli, *ibid.* **52**, 1640 (1995); A. P. Horsfield, A. M. Bratkovsky, D. G. Pettifor, and M. Aoki, *ibid.* **53**, 1656 (1996); E. Hernandez and M. J. Gillan, *ibid.* **51**, 10 157 (1995).

[2] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).

[3] S. Ismail-Beigi and T. Arias, Phys. Rev. Lett. (to be published); S. Goedecker, Phys. Rev. B **58**, 3501 (1998).

[4] I. Daubechies, *Ten Lectures on Wavelets* (SIAM, Philadelphia, 1992).

[5] K. Cho, T. A. Arias, J. O. Joannopoulos, and P. K. Lam, Phys. Rev. Lett. **71**, 1808 (1993).

[6] S. Wei and M. Y. Chou, Phys. Rev. Lett. **76**, 2650 (1996).

[7] S. Goedecker and O. Ivanov, Solid State Commun. **105**, 665 (1998); R. A. Lippert, T. Arias, and A. Edelman, J. Comput. Phys. **140**, 278 (1998); T. Arias, Rev. Mod. Phys. (to be published); S. Goedecker, *Wavelets and their Application for the Solution of Partial Differential Equations* (Presses Polytechnique Universitaires et Romande, Lausanne, Switzerland, 1998).