

Wavelets: A new paradigm for quantum chemical calculations?

Tor Flå and Eirik Fossgård
Department of Mathematics
University of Tromsø
N-9037 Tromsø
Norway

Kenneth Ruud
Department of Chemistry
University of Tromsø
N-9037 Tromsø
Norway

Abstract

We describe the possibilities offered by multiscale theories in quantum chemical calculations, with special emphasis on the use of multiwavelets. Our focus will be on the possibilities offered by wavelets to the problems appearing in quantum chemical modelling of molecular materials and biomolecules. We will also briefly outline the recent advances in wavelet theory in the problem of separation of variables that has made it possible to apply wavelet theory to the study of problems in quantum chemistry. We end the paper with an outlook on the future of wavelet theory as a new and competitive formalism for use in quantum chemical calculations.

1 Introduction

Nanotechnology is emerging as one of the most important fields in which new technological advances are expected to be made in the near future in such diverse scientific fields as computers, optical devices and biologically active nanotechnology-based inventions. The small size of nanotechnological devices requires that these be built from individual molecules connected to various kinds of macroscopic interfaces [1, 2]. Such hybrid systems will be particularly important while the field of nanotechnology matures, as a bridge between the old and the new technology paradigms. New nanosized devices may for instance consist of a series of individual molecules which act as molecular motors, as computer chips, as extended arrays of molecule-based computer memory, as molecular

wires or as molecular magnets. These molecules will in turn be attached to a solid/extended substrate functioning as a basis on which the nanotechnological devices are attached and either made to act as a complete nanotechnological device or as a link to the macroscopically working device. An illustration of such a concept for a highly dense, molecule-based computer memory is shown in Fig. 1.

To successfully develop such devices, a detailed understanding of the behavior of individual molecules and the interaction of these molecules with “macroscopic” interfaces will be of great importance for the scientific and technological progress in the field. Fig. 1 clearly illustrates the different scales that need to be modeled in order to faithfully represent the complete device: An infinite solid substrate (in principle) on which a layer of atoms is fixed (the memory), a periodic structure in two dimensions (the nanotube), as well as a single reaction center (molecule) at the tip of the nanotube that will store and retrieve data by interaction with atoms on the solid layer. To model such a complex system, low-scaling and flexible computational techniques are required.

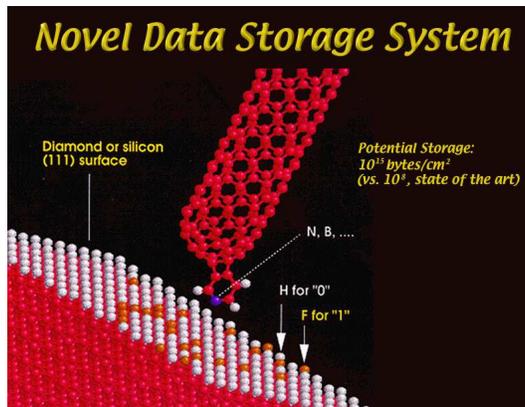
Due to the small size of the molecules expected to be important in nanotechnological devices, as well as the difficulties faced in connecting these molecules to solids, macroscopic surfaces or wires, the cost of production can be very high. Theoretical modeling can therefore be expected to be an integral part of the development process of nanotechnological devices, in particular as a way of reducing the cost of production by exploring a large array of molecules and connecting surfaces/wires to which these molecules can be attached, and in this way preselect the most promising molecules and connecting surfaces/wires, in a manner similar to the methods already in use in connection with drug design.

The problem of modelling molecules at different scales is also present in biomolecular modelling. In many cases one is interested in a detailed understanding of the interaction of an active substrate, such as a specific medical drug, with an active site in for instance a protein. In this specific case, detailed energetics is only required for the drug and the active site of the protein, yet the protein backbone need to be present in order to provide the necessary sterical and energetic stabilization of the interacting complex, though it can be probably be represented in a more approximate manner.

Whereas there are a number of highly successful theoretical approaches available for studying isolated molecules [3, 4, 5] as well as computational methods suited for the study of solid-state systems of either a metallic or molecular crystalline nature [10], theoretical methods suited for studying the interaction of one or more isolated molecules with a solid-based substrate are much less developed. Since the properties of individual molecules can be expected to be significantly affected by the solid surface/interface, the successful modeling of nanotechnologically relevant systems requires a faithful representation of the effects of these interactions, and thus the computational approach must be able to calculate—with the accuracy required—the energies and densities of the various parts of the total system.

From a computational point of view, the system that needs to be modeled becomes very large and complex. One possible approach for dealing with this

Figure 1: A suggested layout for an ultra-dense, molecule-based computer memory. Illustration taken from the NASA Nanotechnology gallery (<http://www.ipt.arc.nasa.gov/gallery.html>).



problem is to realize that energies, as well as variations in the electron density, vary on the different scales present in the system (at the atomic level, within each molecule, in the larger solid substrate, and between the molecules and the solid) and exploit these different energy/density variation scales when solving the problem. Furthermore, not all parts of the system may require as high/dense representation as that part of the system at which reactions or interactions occur. Based on this view of the problem, a multiscale approach appears as an attractive approach to this complex problem. We will limit our discussion here to the use of multiwavelets to attack this problem. Furthermore, the approach described in this paper is restricted to the electronic degrees of freedom. However, with a fast computation of the electronic energies, this approach can be used in combination with ordinary drivers for the calculation of the nuclear motion in for instance molecular dynamics simulations.

2 The quantum mechanical many-particle problem

We do not intend in this contribution to give a complete account of the equations that need to be solved when modelling a molecular system. However, in order to give the reader an qualitative understanding of the mathematical problem at hand, we will briefly introduce here the electronic Schrödinger equation and outline the way in which this equation is being solved in most quantum chemical programs.

In the commonly used Born–Oppenheimer approximation [6, 7] in which the electrons are assumed to be adjusting themselves instantaneously to the position

and movement of the nuclear framework such that the nuclei can be considered to act as a static potential in which the electrons move, the properties of the molecular system is determined by the electron density and electronic energy only. In the very successful Density functional theory (DFT) [5] approach to solving the electronic problem

$$E[\rho] = T[\rho] + V[\rho] + V_{\text{ext}}[\rho], \quad (1)$$

where $T[\rho]$ is the kinetic energy functional arising from the kinetic energy of the electrons, $V[\rho]$ the potential energy functional arising from the interactions between the electrons, and $V_{\text{ext}}[\rho]$ the external potential acting on the electron density (which in most cases is due only to the nuclear framework). This energy functional is exact and it is a local functional of only the three spatial coordinates \mathbf{r} .

However, it has proved difficult to find accurate functionals for the kinetic and potential energy contributions for modelling molecular systems due to the strongly varying electron density in these systems (simple and accurate functionals do exist for metallic systems or systems with a uniform and homogeneous electron density), and the recent success of density functional theory is to a large extent due to the Kohn–Sham approach to DFT [8]. In the Kohn–Sham approach we consider instead of the strongly interacting electron density of a molecule, a system of non-interacting electrons in which the electrons move in a modified effective potential that gives rise the same electron density as the strongly correlated system.

The nonlinear Kohn–Sham equations result from a minimization of the DFT energy functional with respect to variation of occupied Kohn–Sham orbitals $\psi_i(\mathbf{x})$, $i = 1, \dots, M$, which for a closed shell system define the total electron density

$$\rho(\mathbf{r}) = 2 \sum_{i=1}^M |\psi_i(\mathbf{r})|^2.$$

The occupied orbitals are the M lowest eigenfunctions $\{\psi_i, \lambda_i\}$ of the Kohn–Sham operator which implicitly depend on the electron density through an effective potential

$$\begin{aligned} [-\frac{1}{2}\nabla^2 + V_{\text{eff}}(\mathbf{r})]\psi_i(\mathbf{r}) &= \lambda_i\psi_i(\mathbf{r}), \\ V_{\text{eff}}(\mathbf{r}) &= V_e(\mathbf{r}) + V_c(\mathbf{r}) + V_{xc}(\mathbf{r}), \\ V_e(\mathbf{r}) &= -\sum_{\alpha} \frac{Z_{\alpha}}{|\mathbf{r} - \mathbf{r}_{\alpha}|}, \\ V_c(\mathbf{r}) &= \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \\ V_{xc}(\mathbf{r}) &= f[\rho], \end{aligned} \quad (2)$$

where $-1/2\nabla^2$ is the operator for the kinetic energy of the electrons, and the effective potential consists of a nuclear attraction potential $V_e(\mathbf{r})$, a Coulomb

repulsion between the electron densities of two electrons $V_c(\mathbf{r})$, and a remaining, hopefully small, potential accounting for the exchange and correlation interactions among the electrons $V_{xc}(\mathbf{r})$. The form of $V_{xc}(\mathbf{r})$ is however not known, though a few limiting behaviours of the functional has been derived.

Kohn–Sham DFT has proven to be a computationally cost-effective approach for both the theoretical modeling of molecules and for the modeling of extended, periodic systems [9]. Recently, linear-scaling based approaches has gradually been removing the boundaries between these two extremes [10, 11]. In current computational chemistry, the Kohn–Sham orbitals are for molecules in most cases represented in terms of basis sets consisting of Gaussian orbitals. The molecular orbitals $\psi_i(\mathbf{r})$ are written as a linear combination of Gaussian orbitals

$$\psi_i(\mathbf{r}) = \sum_{\mu} C_{i\mu} \chi_{\mu}(\mathbf{r}_K) = \sum_{\mu} C_{i\mu} P_{\mu}(\mathbf{r}_K) \exp(-\alpha_{\mu} r_K^2) \quad (3)$$

where the expansion coefficients $C_{i\mu}$ are referred to as molecular orbital coefficients, and where we have indicated that the electronic coordinates are given relative to the nuclear center K to which the Gaussian basis function is attached. $P_{\mu}(\mathbf{r}_K)$ denotes a Cartesian polynomial $x_K^i y_K^j z_K^k$, in general of order less than 4 ($i + j + k \leq 4$). In principle the atomic basis set should be infinitely large, but is general restricted to about 20-30 functions for each atom in the molecule.

For extended periodic systems, the most convenient approach is the representation in terms of Gaussian plane waves [9, 12] which easily exploits the periodicity of the system and allows for fast evaluation of the molecular integrals:

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{k}} C_{i\mathbf{k}} \exp(i\mathbf{k}r) \quad (4)$$

where \mathbf{k} is a three-dimensional wave vector.

However, both approaches are somewhat inadequate when facing the challenge of modeling a system which can be partitioned into a molecular subsystem and one or more extended or periodic structures.

When starting from the “molecular” point of view one could treat an extended system by exploiting the recent advances made in linear scaling methodology [10, 11], as these methods allow increasingly larger systems to be treated using the DFT approach. Solid states still have to be treated by applying periodic boundary conditions. Periodic boundary conditions have been implemented using Gaussian basis sets [10, 13], recently also in a scheme that avoids the numerical instabilities that may arise from the infinite Coulomb summations when larger, diffuse basis functions are used [11].

However, Gaussian basis functions have three significant drawbacks [14]:

- the integrals are computationally expensive to evaluate;
- the coupling between the basis functions and the nuclei makes the evaluation of nuclear forces expensive, that is, derivatives of the electronic energy with respect to nuclear distortions;

- Gaussian basis functions centered on different nuclei are not orthogonal, potentially yielding numerical instabilities in large molecules and basis sets.

For this reason, Gaussian plane waves have attracted much attention in periodic-boundary problems since the Gaussian plane waves have several advantages [14]:

- electron-electron integrals are almost trivial to evaluate;
- they can be chosen independently of the nuclei, making the calculation of nuclear forces very simple;
- they have a uniform grid of nodal surfaces allowing for the easy calculation and extrapolation of correlation energies.

However, from the viewpoint of chemistry at multiple scales, Gaussian plane waves also have certain disadvantages [14]:

- in order to keep the basis at a manageable size, the Gaussian plane-waves cannot model the atomic core region accurately. This can be circumvented using pseudo-potentials [9] at the expense of significantly higher computational complexity;
- the plane-wave basis is in general very large (roughly 30000 basis functions even for small molecules), without necessarily obtaining sufficient accuracy;
- they cannot easily be applied for isolated molecules. The use of very large periodic boxes minimizes the interactions between the molecules in the different boxes, thus hopefully giving a faithful representation of the isolated system.

Recently, an approach was presented in which the Gaussian plane-wave approach and the Gaussian basis function approach were combined [14].

3 Multiwavelets in the calculation of Kohn–Sham orbitals, densities and energies

The efficient computation of molecular orbitals for a many particle system described by *e.g.* DFT requires that spatially adaptive, fast $O(N)$ methods for multiscale representations are developed and combined with a low separation rank representation of integral operators described in terms of convolution kernels such as the Poisson kernel and bound-state Helmholtz kernel. Until recently it has not been clear how to develop efficient algorithms to represent operators by multiresolution methods in more than one dimension. Using the low separation rank of integral kernel operators combined with multiwavelets, Beylkin *et*

al. (see for instance Refs. [15, 16, 17]). has developed efficient multiresolution algorithms with sparse $O(N)$ representations in more than one dimension. The work is based on an earlier observation that singular operators have a natural representation in a multiresolution basis [18]. It is well known that when thresholded to an accuracy $O(\varepsilon)$, a large class of singular operators (Calderon-Zygmund operators) have a band-structure representation in a multiresolution basis.

Multiwavelets has the same important properties as wavelets, such as vanishing moments, orthogonality and compact support. The basis functions have no overlap on a given scale and are organized in small groups of several functions with the same support. The use of a localized and adaptive basis of multiwavelets can be combined with a separation of functions and operators that is a generalization of the separation of variables. A separated representation of a function or the kernel of an operator in many dimensions to an arbitrary, but finite precision ε is given by a sum of products of functions of lower dimension [15, 16] $f(\mathbf{x}) \approx \sum_{k=1}^r s_k \Phi_k^1(x_1) \dots \Phi_k^d(x_d)$, $|f(\mathbf{x}) - \sum_{k=1}^r s_k \Phi_k^1(x_1) \dots \Phi_k^d(x_d)| \leq \varepsilon$. The set of functions $\{\Phi_k^l(x_l)\}$ can either be fixed or chosen adaptively such that for a fixed ε , one adapts the coefficients in the expansion in order to minimize the separation rank r . Similar low-rank representations exist for vectors, matrices, operators and kernels of operators acting in multiple dimensions in terms of low-rank sums of tensor products of the corresponding sets of objects acting in each dimension separately. For example, the rank r -separable representation of a kernel of an operator in terms of elementary kernels acting in one dimension is given as $K(\mathbf{x}, \mathbf{y}) \cong \sum_{m=1}^r s_m \Phi_m^1(x_1, y_1) \dots \Phi_m^d(x_d, y_d)$ to the accuracy $O(\varepsilon)$.

This scheme appears rather straightforward, but it is at present not well understood how to develop such expansions systematically for operators in general except for in some special cases which include the Green's functions of the Poisson and bound-state Helmholtz operator. A low separation rank expansion for a large class of functions can however be obtained as we point out below by multiwavelets expansions combined with thresholding of the expansions coefficients. It is a nonlinear method to represent a function in a multi-dimensional space while only using a small number of parameters. Many linear algebra operations like matrix and vector additions, matrix-matrix and matrix-vector multiplications, and vector inner products can be performed while keeping all objects in a separated form with low separation rank. Operations in d dimensions can in the separated representation be performed by using tensor product combinations of one-dimensional operations. One can then achieve a computational complexity which formally scales *linearly with rd* . Two practical questions have to be addressed in order to make the above ideas work:

- How does one find a close to optimal but not necessarily unique separation rank for a given function and operator?
- How does one find corresponding robust numerical objects and algorithms which maintain the low separation rank?

The first question can at present not be answered in general, though it has been possible to achieve close to optimal separations for a few special cases. In particular, two methods have been tried: In the first approach the operators, functions of operators or kernels are represented by an integral suitable for low-rank quadrature approximations. In the second approach, a generalized least-square procedure is used to search for low separation rank approximations to matrices acting on functions on a multidimensional grid. We can for instance use Legendre multiwavelets and nonstandard or standard representations for each dimension separately since the operators and functions can be represented in the low-separable rank format. By nonlinear thresholding of multiwavelets coefficients and matrices of functions and operator kernels it is possible to obtain robust numerical representations which can maintain a low separation rank. This will be discussed in the next section. In order to represent kernels of inverses of operators such as the Green's function for the bound-state Helmholtz and Laplace operators, the first approach is particularly well suited. These operators are central in the Kohn–Sham equations represented as a generalized eigenvalue problem as discussed below. We have found representations (unpublished) of these operators by methods similar to those of Hackbusch *et al.* [19, 20] by using an integral representations of these Green's functions integrated over an additional dummy variable on the interval $(0, \infty)$ with Gaussian, separable kernels. This can then be combined with suitable mappings in the dummy variable to obtain discrete sinc-quadratures as an approximation with a low-rank mixture of separable Gaussians with coefficients depending on the quadrature.

4 Multiwavelets and nonlinearly separated representations

It has been shown that singular operators such as Green's functions have a natural representation in a multiresolution basis. In density functional theory and many other modelling problems, the choice of basis should accommodate integral operators, differential operators and also boundary conditions. The multiwavelets developed by Alpert[21] satisfy some of these requirements. The basis functions do not overlap on a given scale and are organized in vector functions of dimension k sharing the same support on a dyadic interval on $[0, 1]$. Because of the vanishing moments of the basis functions, a large class of integro-differential operators has effectively sparse representations. Effectively sparse means a representation which deviates from a sparse matrix only by a small norm. In one dimension a function $f(x)$ on the interval $[0, 1]$ is linearly truncated to the finest scale n by $f(x) \approx f^n(x) = \sum_{l=0}^{2^n-1} s_l^n \cdot \varphi_l^n(x) \in V_n^k$, where the vector space spanned by the k dimensional vector scaling functions $\{\varphi_l^n(x)\}$ has dimension $2^n k$. The index l denote labels of the 2^n grid points at scale n , and below we label by the multiindex $\mathbf{l} = (l_1, l_2, \dots, l_d)$ the 2^{nd} grid points in dimension d . In Fig. 2 we present for the case $k = 6$ both the six scaling functions and the multiwavelets functions at scale $n = 0$ for the Legendre polynomial and

the interpolating polynomial cases. The multiwavelet space W_{n-1}^k is required to be an orthogonal complement of V_{n-1}^k in V_n^k . The multiwavelets basis, which consists of scaling functions at the coarsest scale n_0 and wavelets functions as a basis for the complementary 'difference' wavelets spaces at all scales n' from the coarsest up to the finest $\{\psi_l^{n'}(x)\}$, are required to satisfy orthogonality and polynomial moment conditions. In d dimensions we can get a representation with a low separation rank by introducing the tensor-product scaling space $V_n^{kd} = V_n^{k,1} \otimes \dots \otimes V_n^{k,d}$ with the tensorproduct scaling basis $\{\varphi_1^n(x_1, x_2, \dots, x_d) = \varphi_{l_1}^n(x_1) \otimes \dots \otimes \varphi_{l_d}^n(x_d)\}$ which spans V_n^{kd} and gives rise to a separable representation of functions in d dimensions. This representation will however not necessarily give rise to an optimal rank if it is linearly truncated.

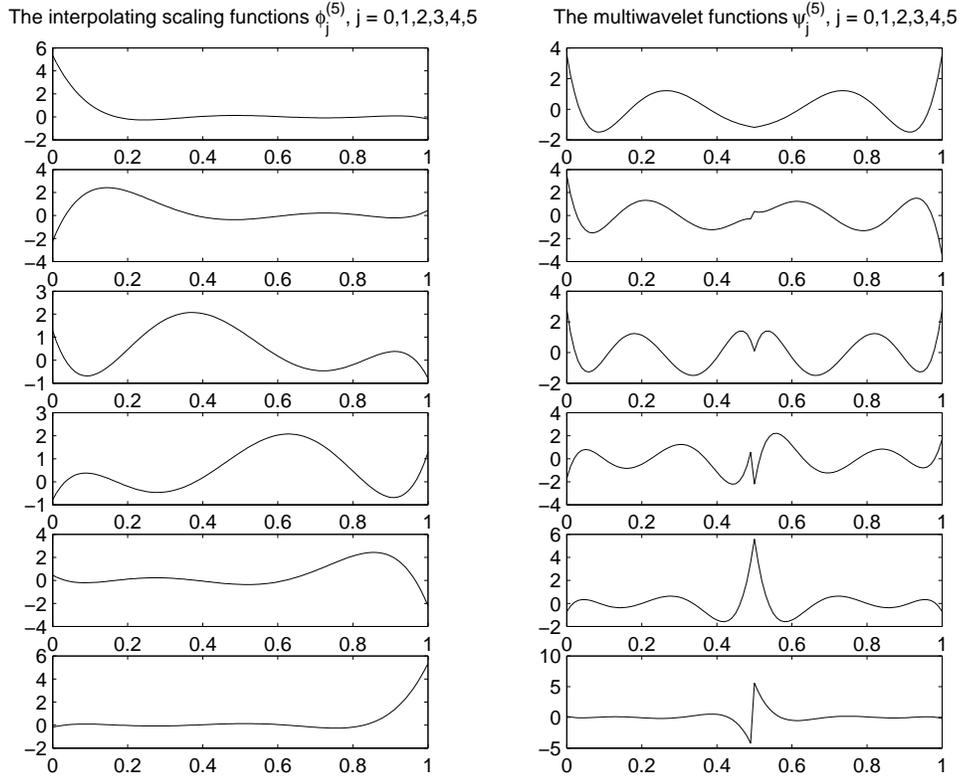


Figure 2: The interpolating scaling functions and the multiwavelet functions of order 5.

In multiple dimensions there are now several ways to decompose the tensor scaling space in terms of sums of tensor products of one-dimensional multiwavelets and scaling spaces. We can again give a linearly truncated expansion of a multi-dimensional function as $f(x_1, \dots, x_d) \approx f_n(x_1, \dots, x_d) = \sum_1 s_1^n \cdot$

$\varphi_1^n(x_1, \dots, x_d)$. In wavelets theory, the separated representation to order ε is found by nonlinear thresholding of the multiwavelets coefficients below a given threshold $\delta^{n'}(\varepsilon)$ adjusted at each scale n' such that the error in each position is kept below $O(\varepsilon)$. The multiwavelets representation is given by repeated application of the orthogonal complement between coarser and coarser scales until the coarsest scale vector space $V_{n_0}^{kd}$ with scale $n_0 < n$ is reached. This gives rise to d -dimensional multiwavelets spaces $W_{n'}^{kd}$ which can be interpreted as difference spaces at all scales between the coarsest up to the finest. However, an important difference is that in d dimensions the wavelets basis will consist of mixed tensor products of both scaling functions and at least one coordinate label with a wavelets function. We index this mixed list by an additional multi-index $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_d) \in \{0, 1\}^d$ excluding $\sigma = \mathbf{0}$ except at the coarsest scale. Here $\sigma = 0$ or 1 signifies a scaling function or a wavelets function for the corresponding coordinate label. The total set of d -dimensional multiresolution basis functions can therefore be labeled as $\{\psi_{\sigma,1}^{n'}\}$.

The multiwavelets coefficients $\{w_{\sigma,1}^{n'}(\varepsilon)\}$ is the list of coefficients truncated if their norm is below a threshold $\delta^{n'}(\varepsilon)$. This is what is meant by an adaptive thresholding of the multiwavelets representation and is one of the major advantages of wavelets. In the multiwavelets representation, the number of terms and scales included will not be fixed, but will be nonlinearly adapted to each grid point. Here dynamic datastructures which can ensure that we obtain the optimal complexity in elementary computations with nonlinearly truncated multiwavelets have to be developed (see Ref. [22]). These data structures will then have to be optimized with respect to the low separation rank representation of fixed operators. A separated representation of the truncated multiwavelets coefficients of different functions simply corresponds to a rearrangement of the labelling of the multiwavelets coefficients by coordinates and letting the indexes vary as in a onedimensional multiwavelet expansion ($w_{(\sigma_1, l_1), \dots, (\sigma_d, l_d)}^{dn'}$) in each coordinate successively ¹. This follows from an alternative multiwavelets expansion ordered by coordinates, and a similar rearrangement of the basis.

For separable operators with kernel $K(\mathbf{x}, \mathbf{y})$ we have to introduce double indexes for translation and labelling of scaling/multiwavelets for each coordinate $\mathbf{l}(2) = (\mathbf{l}, \mathbf{l}')$ and $\sigma(2) = (\sigma, \sigma')$. For a separable operator of low rank, we have an expansion of the kernel in tensor products of elementary kernels acting in one dimension $K(\mathbf{x}, \mathbf{y}) = \sum_{m=1}^r s_m \otimes_{b=1}^d K_{m,b}(x_b, y_b)$. We use a similar method as for functions and find that the separated nonstandard representation in d dimensions can be rearranged in tensor products of one-dimensional nonstandard representations with coefficients

$$r_{NS}^{d(n_0, n-1)} = \sum_{m=1}^r s_m \otimes_{b=1}^d r_{NS, m}^{b, (n_0, n-1)}.$$

¹The indexes in the separated representation vary successively for each spatial direction b as in the above one-dimensional multiwavelets representation, *i.e.* $\{(\sigma_b, l_b, n')\}$ where $\sigma_b \in \{0, 1\}$ if $n' = n_0$ and $\sigma_b = 1$ if $n' \in \{n_0 + 1, \dots, n - 1\}$

The nonstandard representation of operators consists of a series of block-matrices indexed by $\{\sigma(2), \mathbf{l}(2), n'\}$. The separated version of these block-matrices are relabelled such that they successively depend on the nonstandard set of indexes for each elementary kernel, $\{\sigma_b(2) = (\sigma, \sigma'), l_b(2) = (l, l'), n'\}$. The action of these elementary operators in the separated version of the nonstandard representation can be induced by multiple one-dimensional elementary operations. For translational invariant operators with convolution kernels $K(\mathbf{x}, \mathbf{y}) = K(\mathbf{x} - \mathbf{y})$, the representation of the kernels simplifies to discrete convolution kernels depending only on the differences between the positional grid labels, i.e. $\mathbf{l}(2) \rightarrow \mathbf{t} = \mathbf{l} - \mathbf{l}'$, $l_b(2) \rightarrow t_b = l_b - l'_b$. This leads to elementary kernels with matrices at each scale as $r_{m, \sigma_b(2), l_b(2)}^{b, n'} = r_{m, \sigma_b(2), t_b}^{b, n'}$ and that the action becomes of the discrete convolution type. In this case, the basis for the kernel can be replaced by $k \times k$ correlation functions of the scaling functions in the scaling basis representation or by mixed scaling and multiwavelets correlation functions.

The low separation rank tensor product representation of both functions and operators ensures that the action of operators scale essentially as in one dimension. An important advantage of performing computations in a wavelets basis is that the truncated forms of operators in the standard scaling basis and nonstandard multiwavelets representation of low separation rank operators, can be thresholded *separately in each dimension* to sparsely banded matrices. This gives fast algebraic operations in that for instance the product of operators in one dimension can be performed in $O(-(\log \varepsilon)N)$ operations². A remarkable feature of the nonstandard multiwavelets basis representation of operators is that it leads to a block structure for which the different scales are completely uncoupled. Since low separation rank operators already give a sum of tensor product representations where each dimension appear independently of each other, and this opens the way for an efficient representation of the action of operators which is of order $O(r^2N)$.

We can summarize the advantages of multiwavelets and the low separation rank (LSR) approach as:

- Multiwavelets form an orthonormal basis
- The basis functions has disjoint local support
- All basis functions except the scaling function have vanishing moments
- Two-scale relationships are available for computations at adjacent scales
- Thresholding of the multiwavelets coefficients lead to adaptive, sparse, nonlinear approximations with no interaction between the scales for functions³ and operators

² $N = 2^n$ is the size of the finest grid at scale n .

³The point-wise local linear truncation error [23] using Legendre polynomial scaling function basis of order k at finest scale n is for a function f that is at least k times differentiable in one dimension: $\varepsilon_p \leq C 2^{-nk} \frac{2}{4^k k!}$, $C = \sup(f^{(k)})$. A finite singularity at a node of the adaptive mesh give still a high order of convergence. The nonlinear thresholded, adaptive

- The multiwavelets representation for a large class of operators and functions is sparse
- Fast algorithms with guaranteed precision for many operations
- The LSR representation for many operators in d ($d = 3$ or 6 for Density Functional or Hartree–Fock theory, respectively) dimensions implies that operations are effectively one-dimensional given a representation with a low separation rank tensor-product.
- Storage requirements are low
- It removes the curse of dimensionality (Complexity larger than $O(N^d)$) for operators and functions represented with a low separation rank

5 DFT solution scheme

In order to solve the general DFT problem described in Sec. 2 in a multiwavelets approach, we reformulate the eigenvalueproblem to a generalized eigenvalue problem:

$$\begin{aligned} \psi + 2G_\mu * (V\psi) &= \psi + (T - \lambda I)^{-1}(V\psi) = 0; \quad \mu = \sqrt{-2\lambda} \\ (-\nabla^2 + \mu^2)G_\mu(\mathbf{x}, \mathbf{x}') &= \delta(\mathbf{x} - \mathbf{x}') \end{aligned}$$

Here $T = -\frac{1}{2}\nabla^2$ is the kinetic energy operator. The solution of the bound-state Green’s function for the simple case of zero boundary conditions at infinity is

$$G_\mu(\mathbf{r}) = \frac{e^{-\mu|\mathbf{r}|}}{4\pi|\mathbf{x}|}, \quad \mu \geq 0.$$

We immediately realize that we need the following to be able to solve the above generalized eigenvalue problem efficiently:

- We need an iterative procedure to solve the above generalized eigenvalue problem which is nonlinear through the density ρ .
- We need a sparse and adaptive representation of the Poisson kernel operator $P(\cdot) = \int \frac{1}{|\mathbf{x}-\mathbf{x}'|} \cdot d\mathbf{x}'$ and the bound state Green’s function operator $G_\mu(\cdot) = \int \frac{e^{-\mu|\mathbf{x}-\mathbf{x}'|}}{4\pi|\mathbf{x}-\mathbf{x}'|} \cdot d\mathbf{x}'$.
- We need sparse, adaptive representations of the eigenstates $\{\psi_i\}$ and the nonlinear functions of these given by the electron density ρ and the exchange potential $V_{xc} = f(\rho)$.

approximation $f^n(\varepsilon_t)$ satisfies the global error estimate [21, 24] $\|f - f^n(\varepsilon_t)\|_2 \leq \varepsilon_t = \varepsilon \|f\|_2$ if we threshold to zero multiwavelets coefficients $\|w_i^{n'}(\varepsilon_t)\| \leq \delta^{n'}(\varepsilon_t) = 2^{-n'/2}\varepsilon_t$ where ε_t is the desired order of the approximation.

The molecular coordinates are assumed to be in the cube $[-L/2, L/2]^3$ containing the molecule and these are mapped into the unit cube $[0, 1]^3$ by the transformation $x' = x/L + 1/2$ for each coordinate. A zero boundary condition can be enforced by zero-padding the multiwavelets coefficients in the boundary region of the box. The necessary inputs are the nuclear coordinates, the required final precision, and an initial guess for the orbitals. The dimension of the box size (L) can according to Harrison *et al.* be scaled from an estimate of the highest occupied molecular orbital (HOMO) with energy $-\varepsilon_{HOMO}$, the expected asymptotic orbital decay of $\exp(-\sqrt{-2\varepsilon_{HOMO}})$ and the required precision ε . Harrison *et al.* [17] find that an odd order (k) of the multiwavelets related to the precision of $\varepsilon = 10^{2-k}$ seems to be the most efficient choice in their implementation. Odd order multiwavelets have superior precision in the calculation of derivatives. If the order of the wavelets is too low, generally the functions and operators decay less rapidly leading to more terms in the expansion in order to obtain the same accuracy by thresholding. Too high order of the multiwavelets give an excessive computational complexity for the same accuracy ε . Differentiation and construction of the Laplacian can be performed using a periodic central difference approximation related to the nonstandard multiwavelets basis or alternatively the standard scaling basis [24].

The coupled, nonlinear integral equation formulation of the eigenvalue problem be solved using iterative techniques as Krylov-subspace accelerated inexact Newton (KAIN) method similar to the standard DIIS [25]. Each iteration step consist of the substeps

- Optionally computing and diagonalizing the Fock matrix in the space of occupied orbitals
- Computing the density as the sum of squares of the molecular orbitals
- Computing the Coulomb potential by convolution with the separated form of the Poisson kernel by using the truncated multiwavelets representation in each dimension
- Computing the nonlinear exchange potential as a local function of the density and represent it in the multiwavelets representation
- Iteration of the integral equation by *e.g.* the above accelerated technique once for each of the M orbitals. This requires fast multiplication techniques for the potentials and fast convolution techniques with the separated form of the bound-state Helmholtz kernel in the multiwavelets representation.
- Update of the eigenvalues and orbitals
- Orthogonalize the resulting orbitals

Elementary operations such as linear combination of functions and inner products of functions can be directly performed at the multiwavelets coefficient

level taking into account that two functions may have coefficients at different scales which have to be subsequently truncated after the operation to optimally preserve the targeted approximation $O(\varepsilon)$. Multiplication of two reconstructed functions can be performed locally on the finest scale level and then thresholding the multiwavelets coefficients. One can alternatively recur down one level to preserve the accuracy of the approximation, tabulating the functions on the quadrature grid, multiplying the values, project back into the original scaling function basis and finally threshold the multiwavelets coefficients. A similar procedure can be devised for the application of a function g on a thresholded and separated function f . *i.e.* $g(f(x))$ and also for the above action of convolution kernels.

A complete transition to separable representations combined with one-dimensional multiwavelets representation in each direction in DFT possibly combined with convolution in multiple dimensions for the Hartree–Fock exchange potentials or timedependent potentials are facing several challenges which has to be addressed in order to in an optimal way to overcome the curse of dimensionality:

- An important challenge is to find a separable, truncated multiwavelets representation of the nonlinear exchange potential $V_{xc}(\rho)$ since the exchange potential typically is singular for zero electron density.
- Another important challenge is to find and implement a truncated multiwavelets representation that reflect adaptively both truncated potentials and singular Green’s functions and it’s action on truncated multiwavelets representations of all the molecular orbitals.
- One would like a separated representation that has approximate, algorithmic size-extensivity in the sense of a local and hierarchical scale adaptivity of the separation rank to large scale molecular units and subunits. More generally, finer approximations could be used in modules of crucial importance for the molecular system at hand. For large molecules we believe a modular approach is essential to reflect the importance of the different modules for the quantum molecular problem under study.

Acknowledgments

This work has received support from the Norwegian Research Council through a Strategic University Program in Quantum Chemistry (Grant No 154011/420) and a YFF grant (Grant No 162746/V00). The work has also received support from the Norwegian Supercomputer Program.

References

- [1] G. Y. Tseng and J. C. Ellenbogen. *Science*, **294**, 1293, (2001).

- [2] A. Bachtold, P. Hadley, T. Nakanishi, and C. Dekker. *Science*, **2001**, 1317, (2001).
- [3] D. R. Yarkony, editor. World Scientific, (1995).
- [4] T. Helgaker, P. Jørgensen, and J. Olsen. *Molecular Electronic-Structure Theory*. Wiley, Chichester, (2000).
- [5] Parr, R. G. and Yang, W., 'Density-Functional Theory of Atoms and Molecules', (Number 16 in The International Series of monographs on Chemistry, Oxford Univ. Press, New York , 1989).
- [6] M. Born and R. Oppenheimer *Ann. der Physik*, **84**, 457, (1927).
- [7] M. Born and K. Huang *Dynamical Theory of Crystal Lattices*. Clarendon Press, Oxford, (1954).
- [8] W. Kohn and L. J. Sham *Phys. Rev.* **140**, A1133, (1965).
- [9] M. C. Payne, M. P. Teter, D. C. Allen, and J. D. Joannopoulos. *Rev. Mod. Phys.*, **64**, 1045, (1992).
- [10] S. Goedecker. *Rev. Mod. Phys.*, **71**, 1085, (2000).
- [11] K.Ñ. Kudin and G. E. Scuseria. *Phys. Rev. B*, **61**, 16440, (2000).
- [12] R. Car and M. Parrinello. *Phys. Rev. Lett.*, **55**, 2471, (1985).
- [13] R. Dovesi, E. Ermondi, E. Ferrero, C. Pisani, and C. Roetti. *Phys. Rev. B*, **29**, 3591, (1984).
- [14] L. Füstí-Molnar and P. Pulay. *J. Chem. Phys.*, **116**, 7795, (2002).
- [15] Beylkin, G. and Mohlenkamp, M.J. , *Proc. Natl. Acad. Sci. USA* **99**, 10246 (2002) (Univ. of Colorado, APPM preprint nr. 476, Aug. 2001).
- [16] Beylkin, B., and Mohlenkamp, M. J., 'Algorithms for numerical analysis in high dimensions', Univ. Colorado preprint nr. 519, Feb. (2004), to appear in *Siam J. Sci. Comput.*
- [17] Harrison, R. J., Fann, G. I., Yanai, T., Gan, Z., and Beylkin, J. *Chem. Phys.*, **121**, 11587 (2004).
- [18] Beylkin, G., Coifman, R. , and Rohlkin, V., *Comm. Pure Appl. Math.* **44**, 141(1991).
- [19] Hackbusch, W. and Khoromskij, B.N. , Hierarchical Kronecker tensor-product approximation to a class of nonlocal operators in high dimensions, Preprint , Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig 2004.

- [20] Gavriljuk, I. P. , Hackbusch, W., and Khoromskij, B.N., Tensor-product approximation to elliptic and parabolic solution operators in higher dimension, Preprint 83, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig 2004.
- [21] Alpert, B. 'A Class of Bases in L^2 for the sparse representation of Integral Operators', SIAM J. Math. Anal. **24**,No. 1, 246(1993)
- [22] Walden, J., Filter bank methods for hyperbolic PDEs, SIAM J. Numer. Anal. **36**(4), 1183(1999).
- [23] Alpert, B. , 'Sparse Representation of Smooth Linear Operators, PhD Thesis, Yale University, 1990.
- [24] Alpert, B., Beylkin, G., Gines, D., and Vozovoi, L., J. Comput. Phys. **182**, 149(2002).
- [25] Pulay, P. , Fogarasi, G., Pang, F., Boggs, J. E., J. Amer. Chem. Soc. **101**, 2550(1979).