Journal of Mathematical Chemistry 22 (1997) 117–142

Wavelets for electronic structure calculations

M.E. Brewster, George I. Fann and Zhiyun Yang *

*Environmental Molecular Sciences Laboratory, K1-96, Pacific Northwest National Laboratory***, Richland, WA 99352, USA*
E-mail: me_brewster@pnl.gov, gi_fann@pnl.gov, zyang@wavelet.pnl.gov

Received 7 May 1997

Molecular electronic structure calculations have a multi-scale character through the presence of a set of singularities corresponding to atomic nuclei, and thus there exists a potential to improve the efficiency of these calculations using fast wavelet transform techniques. We report on the development of a one dimensional prototype benchmark problem of sufficient complexity to capture the features of 3-D problems that are being solved today in quantum electronics calculations. Theoretical estimates of decay across scales and spatial distribution of wavelet coefficients for the solutions of the 1-D and 3-D problems are derived and verified experimentally. Equivalence in a multi-resolution context of the solutions of the 1-D prototype and the 3-D problem is established.

1. Introduction

The theoretical prediction of the electronic structure of molecules and materials is essential in computational chemistry. The electronic structure of molecules and materials is in principle determined by solution of the Schrödinger equation. For N electrons this amounts to solution of a $3N$ dimensional, second order linear differential equation, which is not practical when N is large. It is estimated that approximately 15–30% of super computer time at the NERSC and NSF supercomputer centers are consumed by electronic structure calculations [9]. Recently, a multi-resolution analysis (MRA) strategy was introduced by Meyer [8] and Mallat [7], which for a wide class of operators leads to sparse structure and permits fast algorithm for application of these operators to functions. Since the electronic structure problems have a multi-scale character through the presence of a set of point singularities corresponding to the atomic nuclei, whose locations are known, MRA has the potential to provide an order of magnitude improvement in computational efficiency and accuracy.

* This research was supported by Associated Western Universities, Inc., under grants DE-FG06-92RL-12451, DE-FG07-93ER-75912 and DE-FG07-94ID-13228 with the U.S. Department of Energy.
** Pacific Northwest Laboratory is operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under contract DE-AC06-76RLO 1830. This project was funded by DOE Program in Mathematics, Information and Computational Sciences.

The problems being solved today in computational chemistry for electronic structure are fundamentally 3-D, or in some cases, 6-D. As a first paper of a series, we are going to set up a 1-D prototype to the 3-D electronic structure problem, on which we can develop methods and algorithms tailored to the electronic structure problem. Ideally, we want the 1-D prototype problem to be "difficult" in the same way as the 3-D problem so that when we expend effort into developing an efficient algorithm for the 1-D problem the insight, and methods obtained from that research are directly applicable to the development of efficient algorithms for the 3-D problem.

Unfortunately, the standard 1-D problem, the Schrödinger equation for atoms in spherical coordinates, is too easy in some respects and too difficult in others. For example, the singularity at the atomic nucleus can be effectively "hidden" in the solution of the 1-D problem by placing the nucleus at the origin of the coordinate system. However, this makes the problem too easy because (1) in the 3-D problem the solution singularity is still evident even if it is placed at a grid point because the second partial derivatives of the solution become unbounded in a neighborhood of the nucleus and this affects the compressibility of the solution, and (2) in a true molecular structure problem with many nuclei it is not possible to place all nuclei on nodes of an evenly spaced rectangular grid, anyway. On the other hand, certain difficulties are encountered in the 1-D equation because the nuclear potential, $1/r$, has a non-integrable singularity whereas this difficulty does not arise in the 3-D case. Thus we have developed a new 1-D equation to use as a prototype.

In this paper, we derive a prototype 1-D equation motivated by the following specifications:

(1) The 1-D prototype should correspond to a physically meaningful situation. That is, the problem should describes an "actual" electronic structure. Therefore we consider transformed versions of the basic atomic structure Schrödinger equation.

(2) We expect equivalence between the 1-D prototype and the 3-D problem with respect to certain multi-resolution analysis (MRA) properties. To characterize equivalence in a multiresolution context, we examine the two main aspects of compression that are necessary for MRA to succeed: decay across scales and spatial distribution of coefficients. This corresponds to a "phase-plane" or frequency/space analysis. In particular, MRA equivalence means for any given tolerance for error:

   (a) an equivalent refinement in multiresolution level is required;

   (b) an equivalent spatial distribution of significant coefficients occurs.

If these equivalence relations hold for the solution and operators in the problem, then an equivalence of the computational complexity of the algorithms to solve the problems will follow. In this paper we verify the equivalence of the solutions; equivalence of the operators is expected because of the derivation method but is not directly verified.

The paper is organized as follows: In section 2 we review the multiwavelet and multiscaling functions used in the approach. In section 3 we set up the 1-D prototype satisfying the characteristics listed above. In section 4 we present the theoretical derivation of the decay across scales and spatial distribution. In section 5 we numerically confirm the theoretical derivation from section 4, for the 1-D function

$$f(x) = \exp\bigl(-a|x - x_0|^{1/3}\bigr),$$

which is the solution of the prototype 1-D problems, as well as for the solution of the 3-D problem

$$f(x, y, z) = \exp\bigl(-a|r|\bigr),$$

where $r = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$, and a is some positive constant. Finally, we smmarize the paper in section 6 and discuss open questions and future work.

2. Multiwavelets

We will use Alpert's multiwavelet [1] for its orthogonality, symmetry and lack of overlap. Here we provide a short description of the multiwavelet and multiscaling functions on $[0, 1]$.
Let the mother scaling functions be

$$\phi_m(x) = \begin{cases} \sqrt{2m-1}\,P_{m-1}(2x-1), & \text{in } [0,1], \text{ for } m = 1, \ldots, M, \\ 0, & \text{elsewhere,} \end{cases} \tag{1}$$

where $P_m(x)$ is the Legendre polynomial of degree m restricted to $[-1, 1]$. The set $\{\phi_i\}$ so defined is orthonormal.
Let the mother wavelet functions $\{\psi_i\}_{i=1,\ldots,M}$ be such that:

- The set $\{\psi_i\}$ is orthonormal.
- The sets $\{\psi_i\}$ and $\{\phi_i\}$ are mutually orthogonal: $\int_0^1 \psi_i(x)\phi_j(x)\,\mathrm{d}x = 0$, for $i, j = 1, \ldots, M$.
- The wavelet ψ_l has $(M + l - 1)$ vanishing moments: $\int_0^1 x^i \psi_l(x)\,\mathrm{d}x = 0$, for $0 \leqslant i \leqslant M + l - 2$, $l = 1, \ldots, M$.

The set $\{\psi_i\}$ so defined is unique up to factors of -1. Notice that the wavelet corresponding to $M = 1$ is simply the Haar wavelet.

Let $\delta_j = 2^j$, where decreasing index j from 0 to $-\infty$ corresponds to grid tending from coarser to finer. Define the multiscaling and multiwavelet functions as

$$\begin{aligned} \phi_i^{j,k}(x) &= \delta_j^{-1/2}\phi_i\left(\frac{x}{\delta_j} - k\right), \\ \psi_i^{j,k}(x) &= \delta_j^{-1/2}\psi_i\left(\frac{x}{\delta_j} - k\right), \end{aligned} \tag{2}$$

where $i = 1, \ldots, M$, $j = 0, \ldots, -\infty$, $k = 0, \ldots, 2^{-j} - 1$. Both the $\phi_i^{j,k}(x)$'s and the $\psi_i^{j,k}(x)$'s have local support in $[k\,\delta_j, (k+1)\,\delta_j]$.

Once we have the multiwavelet basis, we can talk about the wavelet representation of a function. For any function $f(x)$, we denote its wavelet representation at the j-level as f_j:

$$f(x) \simeq f_j = \sum_{k,i} s_i^{j,k} \phi_i^{j,k}, \tag{3}$$

where

$$s_i^{j,k} = \int_{\delta_j k}^{\delta_j (k+1)} \phi_i^{j,k}(x) f(x)\,\mathrm{d}x. \tag{4}$$

It is known that

$$f_{j-1} = f_j + \sum_{k,i} d_i^{j,k} \psi_i^{j,k}, \tag{5}$$

where

$$d_i^{j,k} = \int_{\delta_j k}^{\delta_j (k+1)} \psi_i^{j,k}(x) f(x)\,\mathrm{d}x \tag{6}$$

is the wavelet coefficients.

As we can see, the wavelet coefficients $s_i^{j,k}$ fully determine the approximation of a particular function at the jth level. Knowing $s_i^{j,k}$ at a coarser level and the wavelet coefficients $d_i^{j,k}$ at the same level, we can reconstruct the approximation of that function at a finer level.

In practice, the domain of interests is not necessarily $[0, 1]$. We can always rescale the multiwavelet basis so that they are on the proper interval. In the context of this paper, the domain of interests will be $(-\infty, \infty)$, or numerically $(-a, a)$, where a is sufficiently large. We will always make $\delta_0 = 1$.

In multi-dimensional problems the scaling functions and the multiwavelets are the tensor products of the 1-D multiwavelets.

Let m denote the dimension. Now scalar index i and k in 1-D will be replaced by the corresponding vectors $\mathbf{i} = (i_1, \ldots, i_m)$ and $\mathbf{k} = (k_1, \ldots, k_m)$ in m-D, respectively. For a multiwavelet of order M, the scaling function indices go from 1 to M. The wavelet indices go from $M + 1$ to $2M$. So the index $M + j$ corresponds to the jth wavelet function. For example, $m = 2$, $M = 3$, $\mathbf{k} = (k_1, k_2)$, $\mathbf{i} = (1, 5)$, then

$$c_{\mathbf{i}}^{j,\mathbf{k}} = \int_{[\delta_j k_1, \delta_j (k_1+1)] \times [\delta_j k_2, \delta_j (k_2+1)]} f(x, y) \phi_1^{j,k_1}(x) \psi_2^{j,k_2}(y)\,\mathrm{d}y\,\mathrm{d}x.$$

3. Set up of the 1-D model

In this section we set up a 1-D prototype based on the atomic self-consistent field theory (SCF), which satisfies the features listed in section 1.

Atomic self-consistent field theory [6] approximates the electronic structure associated with a single fixed nucleus and N electrons, with orbitals

$$\phi_{n,l,m}(\mathbf{r}) = R_{n,l}(r)Y_{l,m}(\varphi,\theta),$$

where $Y_{l,m}(\varphi,\theta)$ and $R_{n,l}$ are the spherical harmonics and the radial functions, respectively, $\mathbf{r} = (r,\varphi,\theta)$.

The spherically-symmetric atomic SCF problem is normally stated in terms of the variable $P_n(r) = rR_n(r)$, where $R_n(r) := R_{n,0}(r)$. We will, as representatives, consider the atom with 1 or 2 electrons.

3.1. Hydrogen-like atom: 1-electron

In the case of 1-electron hydrogen atom, the Schrödinger equation is linear and can be written as

$$-\frac{1}{2}P'' - \frac{Z}{r}P = EP, \tag{7}$$

where E is the energy, and Z is a constant.

The exact solution to this equation corresponding to the ground state energy is

$$P(r) = c\,r\,\exp(-\alpha r),$$

where $\alpha > 0$ depends on the energy E of the state, and c is chosen for normalization purposes. In particular, we require that

$$\int_0^\infty P^2(r)\,\mathrm{d}r = 1.$$

3.2. Helium-like atom: 2-electron

In this case we have 2 electrons of opposite spin occupying the same orbital. The interaction between the two electrons leads to a nonlinear problem. To describe this interaction, the equation contains an embedded Poisson problem. However, it is still a scalar equation because the orbitals of the two electrons are the same except for the spin direction.

In this case we have

$$-\frac{1}{2}P'' - \frac{Z}{r}P + \frac{A(r)}{r}P = EP, \tag{8}$$

where

$$-A'' = \frac{1}{r}\,P^2 \tag{9}$$

with the nth eigenstate P_n normalized:

$$\int_0^\infty P_n^2(r)\,\mathrm{d}r = 1. \tag{10}$$

Since (7) can be viewed as a special case of (8) and (9) with $A = 0$, we will therefore consider only the the general form (8) and (9) together with normalization (10). The numerical solution of the 1-electron problem in differential equation form using MRA methods has been considered by Fischer and DeFranceschi [3,4].

In 3-D, the atomic SCF is governed by

$$-\frac{1}{2}\Delta R - \frac{Z}{|\mathbf{r}-\mathbf{r}_0|}R + A(\mathbf{r})R = ER \tag{11}$$

and

$$-\Delta A = R^2, \tag{12}$$

which is an analogue to the 1-D problem, where $\mathbf{r} = (x, y, z)$ and $\mathbf{r}_0 = (x_0, y_0, z_0)$.

3.3. Integral equation formulation

We are going to reformulate (8) and (9) as integral equations. By doing so we gain significant advantages when applying MRA since the representation of the integral operators will be sparse in the multiwavelet space. Also, using the integral formula avoids imposing finite boundary conditions for the free-space problems, and is a natural preconditioning step as well.

We may state the 1-D atomic SCF equations above as integral equations through application of a decaying Helmholtz kernel, with a constant $a > 0$ chosen appropriately. This constant would normally be chosen to reflect the energy of the ground state.

For convenience, we extend the domain of r and s to $(-\infty, \infty)$. Rewrite (8) as

$$-\big(P'' - a^2 P\big) = 2\left(\frac{Z}{|r|} - \frac{A(r)}{|r|} + E + \frac{a^2}{2}\right)P. \tag{13}$$

It is easy to check that the fundamental solution of $-(P'' - a^2 P) = \delta(r)$ is the Helmholtz kernel

$$K_a(r) = \frac{1}{2a}\,\exp\big(-a|r|\big).$$

Therefore, equation (13) (and hence the atomic SCF equation (8)) can be rewritten as a singular integral generalized eigenvalue problem on the real line:

$$P(r) = 2\int_{-\infty}^{\infty} K_a(r-s)\left(\frac{Z}{|s|} - \widetilde{A}(s) + \widetilde{E}\right)P(s)\,\mathrm{d}s, \tag{14}$$

where

$$\widetilde{A}(r) = \frac{A(r)}{|r|}, \qquad \widetilde{E} = \left(E + \frac{a^2}{2}\right)$$

with the nth eigenstate having the symmetry properties

$$P_n(-r) = (-1)^{n-1} P_n(r)$$

and

$$\int_{-\infty}^{\infty} P_n^2(r)\,\mathrm{d}r = 2.$$

Next we consider (9); let $A(r) = rB(r)$, then (9) is equivalent to

$$-\left(r^2 B'\right)' = P^2(r). \tag{15}$$

We want to find the fundamental solution $G(r, s)$ of (15) possessing the following properties:

(1) $-(r^2 G')' = \delta(r - s)$.

(2) G goes to 0 as r goes to ∞.

(3) G is bounded as r goes to 0.

(4) G is continuous at $r = s$.

It is easy to check that

$$G(s, r) = \begin{cases} 1/s, & \text{if } 0 \leqslant r \leqslant s, \\ 1/r, & \text{if } r > s \geqslant 0, \end{cases} \tag{16}$$

is one such a solution. Therefore,

$$B(r) = \int_0^{\infty} G(r, s) P^2(s)\,\mathrm{d}s \quad \text{or} \quad A(r) = r \int_0^{\infty} G(r, s) P^2(s)\,\mathrm{d}s.$$

Again, extend the domains of r and s to $(-\infty, \infty)$, we have

$$\widetilde{A}(r) = \frac{A(r)}{|r|} = \int_{-\infty}^{\infty} G(r, s) P^2(s)\,\mathrm{d}s, \tag{17}$$

where

$$G(s, r) = \begin{cases} 1/|s|, & \text{if } (s/r) \geqslant 1, \\ 1/|r|, & \text{if } 0 \leqslant (s/r) < 1, \\ 0, & \text{if } rs < 0. \end{cases} \tag{18}$$

3.4. Transformations to obtain 1-D prototype

Recall that we expect our 1-D prototype to be equivalent to the 3-D problem with respect to essential MRA properties. For that purpose we make the following change of variables $x = r^3$ and $y = s^3$. Also we would like to change the unknown variable back to R (recall that $R(r) = (1/r)P(r)$).

With these changes, integral (14) is now

$$R(x) = \frac{2}{3}\int_{-\infty}^{\infty} (xy)^{-1/3} K_a\left(x^{1/3} - y^{1/3}\right)\left(\widetilde{E} + \frac{Z}{|y|^{1/3}} - \widetilde{A}(y)\right) R(y)\,\mathrm{d}y,$$

and integral (17) is now

$$\widetilde{A}(x) = \int_{-\infty}^{\infty} v(x,y) R^2(y)\,\mathrm{d}y,$$

where

$$v(x,y) = \begin{cases} (1/3)|y|^{-1/3}, & (y/x) > 1, \\ (1/3)|x|^{-1/3}, & 0 < (y/x) < 1, \\ 0, & xy \leqslant 0. \end{cases}$$

Although of a complicated looking form analytically, this 1-D prototype does indeed have most of the features we are looking for in a benchmark problem. The problem has the following similarities to the 3-D electronic structure problem:

(1) Dilation in the 1-D variable x is equivalent to volume dilation in 3-D by the same factor.

(2) All singularities are integrable.

(3) The problem is self-adjoint, and the associated inner product is unweighted.

One disadvantage of the 1-D prototype problem is the loss of the property of translation invariance that is present in the 3-D problem. This is caused by the use of spherical coordinates to create the 1-D atomic SCF problem, and cannot be avoided.

3.5. Integral representation of the 3-D problem

Analogous to the 1-D formula, we can derive the integral representation of the 3-D problem governed by (11) and (12):

$$R(\mathbf{r}) = 2\iiint_{-\infty}^{\infty} K_a(\mathbf{r}-\mathbf{s})\left(\frac{Z}{|\mathbf{s}-\mathbf{r}_0|} - A(\mathbf{s}) + \widetilde{E}\right) R(\mathbf{s})\,\mathrm{d}\mathbf{s}, \tag{19}$$

where $K_a(\mathbf{r}) = \exp(-a|\mathbf{r}|)/|\mathbf{r}|$ and $\widetilde{E} = E + a^2/2$, and

$$A = \iiint_{-\infty}^{\infty} G(\mathbf{s},\mathbf{r}) R^2(\mathbf{s})\,\mathrm{d}s, \tag{20}$$

where $G(\mathbf{s},\mathbf{r}) = 1/(4\pi|\mathbf{r}-\mathbf{s}|)$.

4. Efficiency of the compressed form

We can take full advantage of MRA only if the functions we are going to work with can be compressed efficiently. Therefore, before we apply MRA to the Schrödinger equation, a full understanding of the compressibility of the operator, the kernel and the solution is necessary.

First let's introduce the notations we are going to use hereafter.

- Let $\widehat{f}(x)$ be an approximation of $f(x)$, we define *error* ε as the L_2 norm of the difference:
$$\varepsilon = \left\| f(x) - \widehat{f}(x) \right\|_2.$$
- In implementation of the wavelet representation of functions, a *cutoff* σ is specified. Wavelet coefficients whose absolute values are less than the cutoff are ignored.

The wavelet coefficients whose absolute values are greater than or equal to the cutoff are called the *significant coefficients* and the wavelet approximation of a function with only significant coefficients is called the *compressed* representation.

As we can see, the efficiency of the compressed representation in the wavelet basis is determined by the total number of significant coefficients over all levels. Hence the number of levels affects the efficiency. This is determined primarily by the rate at which the magnitude of the wavelet coefficients decreases as j or δ_j decreases. This rate is called the *rate of decay across scales* for a particular function. It depends on, and is a good indicator of, the smoothness of the function [5]. The decay is fast if the function is smooth. The decay rate is also affected by the order of the wavelet basis we choose to use.

A particular 1-D function we want to investigate is
$$\mathcal{F}(r) = \exp\bigl(-|r|\bigr) = \exp\bigl(-|x|^{1/3}\bigr),$$
or more generally,
$$\mathcal{F}(x) = \exp\bigl(-|x - x_0|^{1/3}\bigr), \tag{21}$$
since both the kernel and the solution of the 1-D prototype are essentially in this form. We include a shift of x_0 to cover the general situation. Note that this function is continuous, but its derivatives become unbounded at x_0. Although with singularity, this function as well as its derivatives are integrable.

In this section we set the goal of answering the following questions about compressibility:

- What is the finest level J needed for a specified cutoff σ, or at what level J can we be sure that all $d_i^{j,k}$'s below J are negligible, namely, $|d_i^{j,k}| < \sigma$ for all $j < J$?
- What is the total number of the significant coefficients from the coarsest level to level J? If the function has a singularity, then how many of the significant coefficients are near the singularity and how many are away from the singularity?

We will also estimate the error in terms of the cutoff σ.

4.1. Estimates of wavelet coefficients

To see the rate of decay across scales for a function, let's estimate the wavelet coefficients $d_i^{j,k}$ defined by (6). We list together the result for $s_i^{j,k}$ defined by (4) although we do not need it for the decay estimate.

Let's first consider functions without singularities.

Lemma 1. Let $f(x)$ be a function with continuous derivatives everywhere up to order $p_1 - 1$, and a bounded p_1th derivative, where p_1 is a positive integer. Let $p_2 = \min(i-1, p_1)$, $p_3 = \min(i + M - 1, p_1)$. Then

$$s_i^{j,k} = \mathrm{O}\big(\delta_j^{p_2+1/2}\big), \quad d_i^{j,k} = \mathrm{O}\big(\delta_j^{p_3+1/2}\big) \quad \text{as } j \to -\infty.$$

Proof. By definition,

$$s_i^{j,k} = \int_{\delta_j k}^{\delta_j(k+1)} f(x)\phi_i^{j,k}(x)\,\mathrm{d}x = \delta_j^{-1/2} \int_{\delta_j k}^{\delta_j(k+1)} f(x)\phi_i\left(\frac{x}{\delta_j} - k\right)\mathrm{d}x.$$

Making a change of variable $s = x/\delta_j - k$, we have

$$s_i^{j,k} = \delta_j^{1/2} \int_0^1 f(\delta_j k + \delta_j s)\phi_i(s)\,\mathrm{d}s.$$

From the Taylor expansion for $f(\delta_j k + \delta_j s)$ about $x_k = \delta_j k$, we obtain

$$\begin{aligned} s_i^{j,k} = \delta_j^{1/2} \int_0^1 \Big[& f(x_k) + \delta_j s f'(x_k) + \frac{1}{2}\delta_j^2 s^2 f''(x_k) + \cdots \\ & + \frac{(\delta_j s)^{p_1-1}}{(p_1-1)!} f^{(p_1-1)}(x_k) + \frac{(\delta_j s)^{p_1}}{p_1!} f^{(p_1)}(\xi)\Big]\phi_i(s)\,\mathrm{d}s, \end{aligned} \tag{22}$$

where ξ is some value between $k\delta_j$ and $(k+1)\delta_j$. If $i = 1$, $s_i^{j,k} = \mathrm{O}(\delta_j^{1/2})$ from (22). If $i \geqslant 2$, then ϕ_i has $i - 1$ vanishing moments; that is, $\int_0^1 s^j \phi_i(s)\,\mathrm{d}s = 0$ for $j = 0, 1, \ldots, i-2$, and the conclusion follows immediately from (22). Similarly, we can prove the result for $d_i^{j,k}$. The only difference is that ψ_i has $(i + M - 1)$ vanishing moments. □

Corollary 2. Let $f(x)$ be an arbitrarily smooth function, then

$$s_i^{j,k} = \mathrm{O}\big(\delta_j^{i-1/2}\big), \quad d_i^{j,k} = \mathrm{O}\big(\delta_j^{i+M-1/2}\big) \quad \text{as } j \to -\infty.$$

Next let's consider the function $\mathcal{F}(x)$ defined in (21). We will find that those wavelet coefficients whose corresponding wavelet support contains the singularity decay more slowly than those who do not.

Notice that, locally about x_0, $\mathcal{F}(x)$ can be expanded as

$$\begin{aligned}\mathcal{F}(x) &= \exp\bigl(-|x-x_0|^{1/3}\bigr) = 1 - |x-x_0|^{1/3} + |x-x_0|^{2/3} + \cdots \\ &= 1 + \bigl(-\operatorname{sgn}(x-x_0) + \tfrac{1}{2}(x-x_0)^{1/3} + \cdots\bigr)(x-x_0)^{1/3}.\end{aligned}$$

So $\mathcal{F}(x)$ has a Frobenius-type singularity of power $1/3$.

Lemma 3. Let $f(x)$ be a function with a Frobenius-type singularity of power α, i.e., locally about x_0, $f(x) = f_0(x) + f_1(x)(x-x_0)^\alpha$, where $f_0(x)$ is arbitrarily smooth, and $f_1(x)$ is bounded locally at x_0. Then in the support containing x_0,

$$s_i^{j,k} = \mathrm{O}\bigl(\delta_j^{\min(i-1/2,\alpha+1/2)}\bigr) \quad \text{and} \quad d_i^{j,k} = \mathrm{O}\bigl(\delta_j^{\min(i+M-1/2,\alpha+1/2)}\bigr).$$

Proof. Following the same steps in lemma 1 we have

$$\begin{aligned}s_i^{j,k} &= \delta_j^{1/2}\int_0^1 f(\delta_j k + \delta_j s)\phi_i(x)\,\mathrm{d}s \\ &= \delta_j^{1/2}\biggl(\int_0^1 f_0(\delta_j k + \delta_j s)\phi_i(s)\,\mathrm{d}s + \int_0^1 f_1(\delta_j k + \delta_j s)(\delta_j k + \delta_j s - x_0)^\alpha \phi_i(x)\biggr).\end{aligned}$$

If $s_i^{j,k}$ has a support containing x_0, then

$$|\delta_j k + \delta_j s - x_0| < \delta_j.$$

Hence, from corollary 2,

$$s_i^{j,k} = \mathrm{O}\bigl(\delta_j^{i-1/2}\bigr) + \mathrm{O}\bigl(\delta_j^{\alpha+1/2}\bigr).$$

Similarly,

$$d_i^{j,k} = \mathrm{O}\bigl(\delta_j^{i+M-1/2}\bigr) + \mathrm{O}\bigl(\delta_j^{\alpha+1/2}\bigr). \qquad \square$$

Corollary 4. For $\mathcal{F}(x)$, the same results in lemma 1 hold for wavelet coefficients whose corresponding wavelet supports do not contain x_0. For wavelet coefficients whose corresponding wavelet support contains x_0,

$$s_i^{j,k} = \mathrm{O}\bigl(\delta_j^{\min(i-1/2,5/6)}\bigr)$$

and

$$d_i^{j,k} = \mathrm{O}\bigl(\delta_j^{5/6}\bigr) \quad \text{as } j \to -\infty. \tag{23}$$

Analogously, we can prove the results for the multi-dimensional situation. We therefore omit the proof here but give the general results.

Theorem 5. Let m denote the dimension of the problem, $c_{\mathbf{i}}^{j,\mathbf{k}}$ denote the **i**th tensor product coefficient at level j, position **k** as defined in section 2, where $\mathbf{i} = (i_1, \ldots, i_m)$, $i_{m'} = 1, 2, \ldots, 2M$ for $m' = 1, 2, \ldots, m$, and $\mathbf{k} = (k_1, \ldots, k_m)$, $k_{m'} = 0, 1, \ldots, 2^{-j} - 1$ for $m' = 1, 2, \ldots, m$. $f(x_1, \ldots, x_m)$ has continuous derivatives of order $(q_{m'} - 1)$ with respect to $x_{m'}$ and bounded derivatives of order $q_{m'}$ everywhere, then

$$c_{\mathbf{i}}^{j,\mathbf{k}} = \mathrm{O}\big(\delta_j^p\big) \quad \text{as } j \to -\infty,$$

where

$$p = \sum_{m'=1}^{m} \min(i_{m'} - 1, q_{m'}) + m/2.$$

As a corollary, if f is smooth enough such that $\min_{m'} q_{m'} \geqslant 2M$, then

$$p = \sum_{m'=1}^{m} (i_{m'} - 1) + m/2.$$

A tensor product wavelet (as opposed to a scaling wavelet) must have at least one index $i_{m'} > 1$. On the other hand there is always a wavelet basis with only one index greater than 1. Therefore for the 3-D $\mathcal{F}(x)$ we are interested in, since $q_{m'} = 1$ for all m's, the slowest decay of the wavelet coefficients is $\mathrm{O}(\delta^{1+3/2})$, i.e., $\mathrm{O}(\delta^{5/2})$.

4.2. Number of levels and number of significant coefficients

Now we are ready to answer the two questions we asked ourselves at the beginning of this section: Under what level J are $d_i^{j,k}$ negligible, and what is the total number of the significant coefficients?

We rewrite the asymptotic relationship (23) as

$$\big|d_i^{j,k}\big| \leqslant C\delta_j^{5/6},$$

where C is a constant. Solving $C\delta_J^{5/6} \leqslant \sigma$ gives

$$J \leqslant \frac{6}{5}\log_2\Big(\frac{\sigma}{C}\Big) \sim \frac{6}{5}\log_2(\sigma). \tag{24}$$

To obtain the asymptotic number of significant coefficients, let's first find out L such that $|d_i^{j,k}| < \sigma$ when $|x - x_0| > L$. It is sufficient that $\mathcal{F}(x) < \sigma$. Without lose of generality, let $x_0 = 0$.

$$\exp\big(-|x|^{1/3}\big) < \sigma \quad \Rightarrow \quad |x| > \big(-\ln(\sigma)\big)^3.$$

Namely, $L = (|\ln\sigma|)^3$. The number of significant coefficients on level 0 is therefore $\mathrm{O}(L)$ (recall that $\delta_0 = 1$).

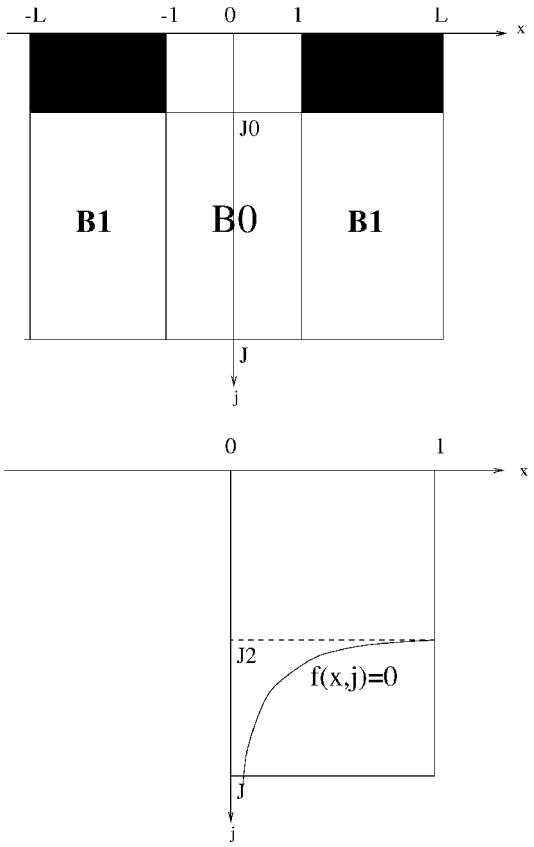

Figure 1. The figure on the top shows the box B. Outside B_0 the significant coefficients lay only in the dark marked regions of B_1. The figure on the bottom depicts the curve $f(x, j) = 0$ cutting through B_0 under which the coefficients are negligible.

Now we know that all significant coefficients take place in the box B: $[-L, L] \times [J, 0]$. We divide this box into two parts: $|x| \leqslant 1$ and $|x| \geqslant 1$. Denote the number of significant coefficients in B_0: $[-1, 1] \times [J, 0]$ as N_0, and the number of significant coefficients in B_1: $B \setminus B_0$ as N_1 (figure 1 shows different regions). Then

$$N_0 \leqslant \mathrm{O}\big(2^{-J}\big) = \mathrm{O}\big(\sigma^{-6/5}\big). \tag{25}$$

In B_1, $\mathcal{F}(x)$ is arbitrarily smooth, so there should be a level $J_1 > J$ below which the coefficients are less than the cutoff. From corollary 2, the slowest decay is $\mathrm{O}(\delta_j^{M+1/2})$ when $i = 1$. Similar to how we determined J, we first rewrite the the asymptotic relationship as

$$\big|d_i^{j,k}\big| \leqslant C' \delta_j^{M+1/2}.$$

Solving $C'\delta_{J_1}^{M+1/2} = \sigma$ gives

$$J_1 = \frac{\log_2(\sigma/C')}{M+1/2} \sim \frac{2}{2M+1}\log_2\sigma.$$

Now, that we have J_1, we know that

$$N_1 = \mathrm{O}\big(L\,2^{-J_1}\big) = \mathrm{O}\big(|\ln\sigma|^3\sigma^{-2/(2M+1)}\big).$$

The asymptotic estimate (25) can be tightened since away from the only singularity point, the function is still arbitrarily smooth although the derivatives are not uniformly bounded, hence the coefficient decay across scales are dramatic.

Lemma 6. For $\mathcal{F}(x)$, the number of significant coefficients in $|x - x_0| \leqslant 1$ is $N_0 = \mathrm{O}(\sigma^{-1/(M+1/2)})$.

Proof. The idea is to find a curve $f(x, j) = 0$ such that in B_0, $|d_i^{j,k}| < \sigma$ under the curve. By doing so we shrink the possible area where significant coefficients could take place. Due to symmetry, we consider only $x \geqslant 0$.

Consider the region in B_0 one grid away from the singularity. One can check that $|d_i^{j,k}| < \sigma$ under the curve in $k - j$ space determined by

$$(\delta_j)^{(M+1/2)} f^{(M)}(k\delta_j) = \sigma, \tag{26}$$

following a similar proof for lemma 1.

This curve decrease monotonically (see figure 1). Denote the point where the curve cuts the right edge of B_0 as $(1, J_2)$. Then $J_2 = J_1$ and

$$J_2 \sim \frac{1}{M+1/2}\log_2(\sigma).$$

Hence the number of significant coefficients in B_0 between level J_2 and level 0 is $\mathrm{O}(\sigma^{-1/(M+1/2)})$. The number of significant coefficients in B_0 between level J and level J_2 can be found to be also $\mathrm{O}(\sigma^{-1/(M+1/2)})$:

$$f^M(k\delta_j) = \mathrm{O}\big((k\delta_j)^{1/3-M}\big).$$

From (26),

$$k_j = \mathrm{O}\big(2^{-5j/2(1-3M)}\sigma^{3/(1-3M)}\big).$$

Hence the number of significant coefficients is

$$\sum_{j=J}^{J_2} k_j = \mathrm{O}(k_{J_2}) = \mathrm{O}\big(\sigma^{-1/(M+1/2)}\big)$$

since $M \geqslant 1$.

The number of significant coefficients within a grid of singularity from level J to level 0 is simply $\mathrm{O}((6/5)\log_2(\sigma))$. Since $M \geqslant 1$, $\mathrm{O}(\sigma^{-1/(M+1/2)})$ dominates. □

Table 1
Compressibility: M_1 and M_3 are used to distinguish M in 1-D and 3-D.

	1-D prototype	3-D SCF
Levels needed (J)	$\sim\frac{6}{5}\lvert\log_2\sigma\rvert$	$\sim\frac{2}{5}\lvert\log_2\sigma\rvert$
Significant wavelet coefficients in $\lvert x-x_0\rvert\leqslant 1$	$\mathrm{O}\big(\sigma^{-2/(2M_1+1)}\big)$	$\mathrm{O}\big(\sigma^{-6/(2M_3+3)}\big)$
Significant wavelet coefficients in $\lvert x-x_0\rvert\geqslant 1$	$\mathrm{O}\big(\lvert\ln\sigma\rvert^3\sigma^{-2/(2M_1+1)}\big)$	$\mathrm{O}\big(\lvert\ln\sigma\rvert^3\sigma^{-6/(2M_3+3)}\big)$
Total number of coefficients in J levels with $\lvert f(x)\rvert\geqslant\sigma$	$\mathrm{O}\big(\lvert\ln\sigma\rvert^3\sigma^{-6/5}\big)$	$\mathrm{O}\big(\lvert\ln\sigma\rvert^3\sigma^{-6/5}\big)$

Similarly one can prove a corresponding result for the 3-D case.

To see the compression, we notice that the total coefficients in B is O($L2^{-J}$), namely, ((ln $|\sigma|$)$^3\sigma^{-6/5}$). The greater the M, the higher the compression. Table 1 summarizes the theoretical results.

We can see from the table that our 1-D prototype represents the 3-D situation well in the sense that the decaying behavior in 1-D and 3-D are about the same when $M_3 = 3M_1$.

4.3. *Error estimate*

In this section we estimate the error of $\mathcal{F}(x)$ at the Jth level. From lemma 4, there exists constants C and C' such that

$$\big|d_i^{j,k}\big|\leqslant C\delta_j^{\alpha}$$

for $d_i^{j,k}$ whose support contains x_0, and

$$\big|d_i^{j,k}\big|\leqslant C'\delta_j^{\alpha'}$$

for $d_i^{j,k}$'s whose supports do not contain x_0, where $\alpha = 5/6$, $\alpha' = \alpha'(i, M) = p_3+1/2$. The smallest α' is hence $3/2$.

Lemma 7. For the function $\mathcal{F}(x)$, the error of its wavelet representation on the Jth level satisfies

$$\varepsilon = \mathrm{O}\big(\delta_J^{\alpha}\big).$$

Proof. From the definition and (5), we have

$$\varepsilon^2 = \big\|f(x)-f_J\big\|_2^2 = \sum_{i=1}^{M}\sum_{j=J}^{-\infty}\sum_{k=0}^{-1+2^{-j}}\big|d_i^{j,k}\big|^2$$

$$\leqslant M\sum_{j=J}^{-\infty}\left(C^2\delta_j^{2\alpha}+2^{-j}C'^2\left(\delta_j^{3/2}\right)^2\right)=M\sum_{j=J}^{-\infty}\left(C^2\delta_j^{2\alpha}+C'^2\delta_j^2\right)$$

$$=M\left(C^2\frac{\delta_J^{2\alpha}}{1-(1/2)^{2\alpha}}+\frac{4}{3}C'^2\delta_J^2\right). \tag{27}$$

Since $2\alpha<2$, the first term dominates, therefore $\varepsilon=\mathrm{O}(\delta_J^{\alpha})$. □

Equation (27) gives the estimate of the error. If we substitute into (27) the formula for J given in (24), we can find the relationship between the cutoff and the error. A rough estimation suggests that $\varepsilon\leqslant\widetilde{C}\sigma$, where $\widetilde{C}$ is a constant depending on M, C and C'. Hence we can on the other hand choose the appropriate cutoff if we want to keep the error within a prescribed tolerance.

5. Numerical results

5.1. The one-dimensional case

In this section, we will study the coefficient decay across scales numerically, for the function $\mathcal{F}(x)=\exp(-|x-x_0|^{1/3})$, make useful observations and pick up some loose ends on motivation in previous discussions. Recall that $\delta_0=1$.

Figures 2–5 show ($\max_{i,k}|d_i^{j,k}|$) v.s. level ($-j$) for $M=1,2,3$ and 4, respectively. Different x_0's are used.

From the plots we can see the decay across scales clearly. We can see that the decay is not monotonic when the singularity, namely x_0, does not concur with a grid point. Nevertheless, the overall behavior is the same for different x_0 no matter if it is a grid point. Notice that this is not true if we do not make the change of variable $x=r^3$. Figure 6 shows ($\max_{i,k}|d_i^{j,k}|$) v.s. level ($-j$) for $M=2$ for the function $f(r)=f(x)=\exp(-|x-x_0|)$, where $x_0=0$, $1/7$ and $1/(64\times 7)$, respectively. We can see that the curve with nondyatic x_0 decays significantly slower than the one with grid point x_0. The reason is that when x_0 is a grid point, the integral over the support of a wavelet (hence the wavelet transform) will not realize its existence, therefore it appears no different than a smooth function. However, when x_0 is not at a grid point, which will be the general case for application to large molecules, the integral over the local support will see that nonsmoothness, and the decay of coefficients across scales will turn out to be slower than a smooth function. That is why we choose to study the test function with arbitrary x_0 and to extend the domain to $(-\infty,\infty)$ in the first place.

To capture the essential behavior of the singularity is one of the motivations why we make the change of variable $x=r^3$ which is the volume in 3-D. Consider the situation when x_0 is very close to 0. Before the change of variable, $\mathcal{F}(r)=\exp(-|r|)$ as well as all its derivatives with respect to r are well bounded by 1, and when the grid is not fine enough, it can be easily overlooked. We could therefore stop at a coarse

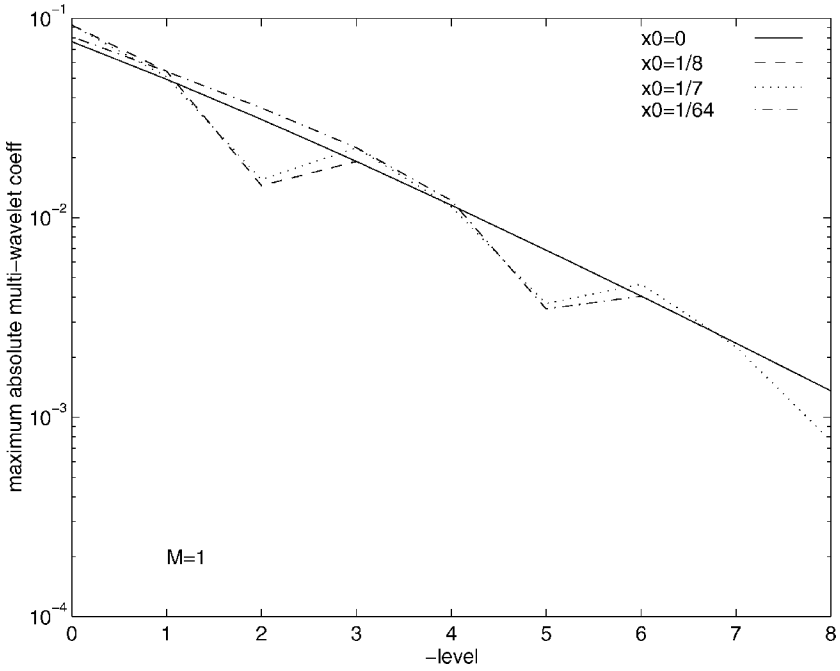

Figure 2. Decay across scales for $M = 1$ (1-D).

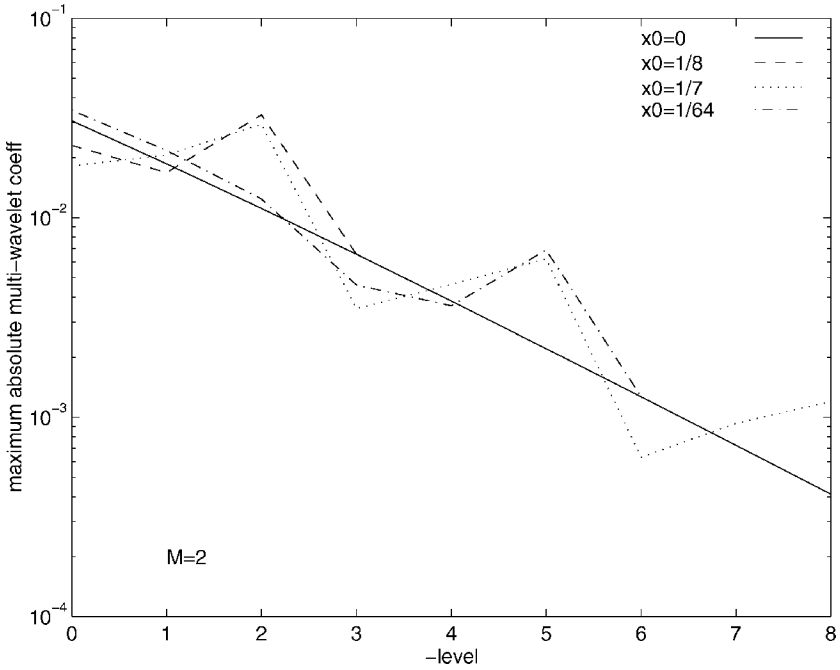

Figure 3. Decay across scales for $M = 2$ (1-D).

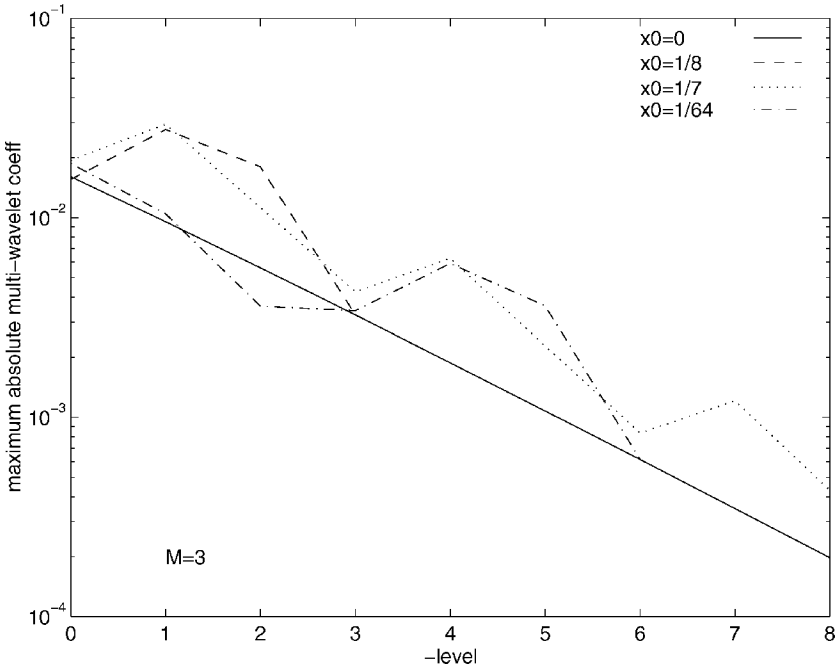

Figure 4. Decay across scales for $M = 3$ (1-D).

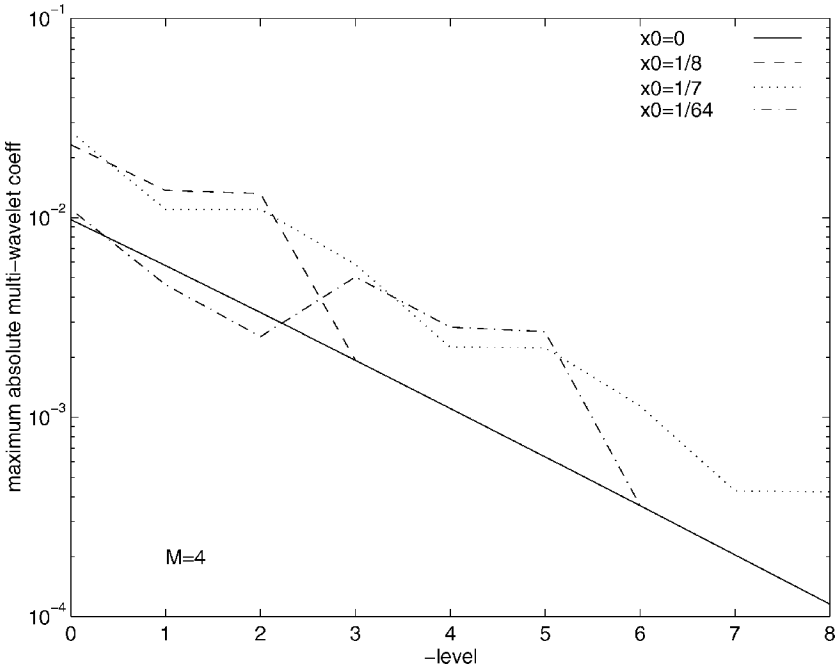

Figure 5. Decay across scales for $M = 4$ (1-D).

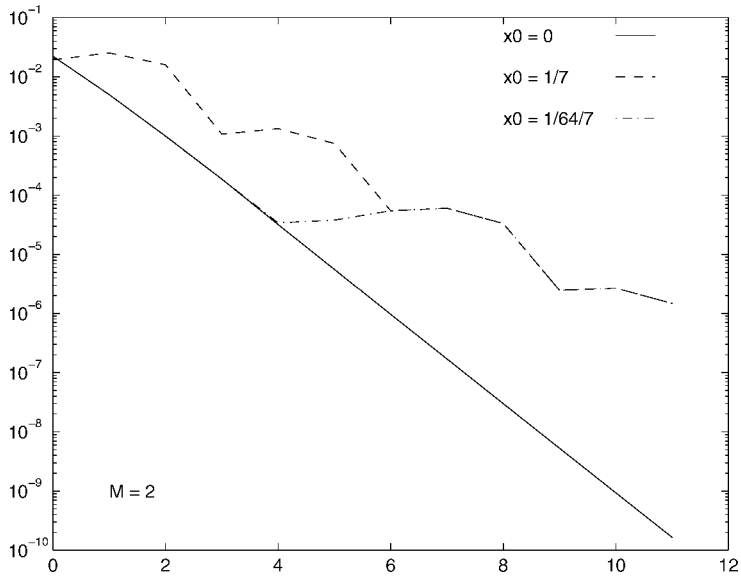

Figure 6. Decay across scales for $f(x) = \exp(-|x - x_0|)$, $M = 2$, $x_0 = 0$, $1/7$ and $1/(64 \times 7)$, respectively.

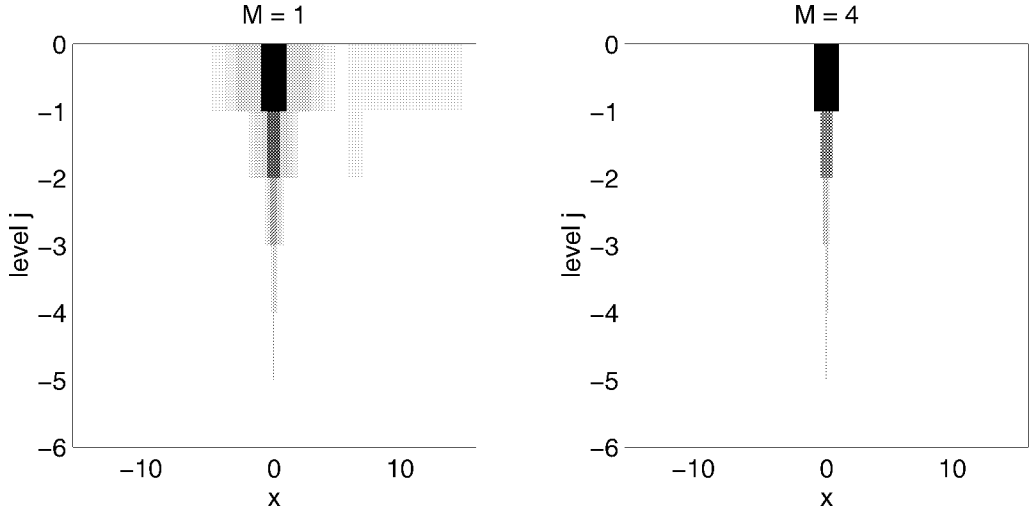

Figure 7. Wavelet coefficient distribution (1-D): The darker the color, the greater the absolute value of the coefficient. $i = 1$, $x_0 = 0$, the order of the multiwavelet is $M = 1$ and $M = 4$, respectively.

level before we capture enough detail to accurately describe the singularity. This is demonstrated by the curve with $x_0 = 1/(64 \times 7)$ in figure 6: The singularity is not captured until the level is fine enough. The change of variable makes the derivative of the function blow up at the singular point, hence the singularity is harder to miss. This is also a better model of the 3-D problem since in 3-D, the derivatives of $\mathcal{F}(r)$

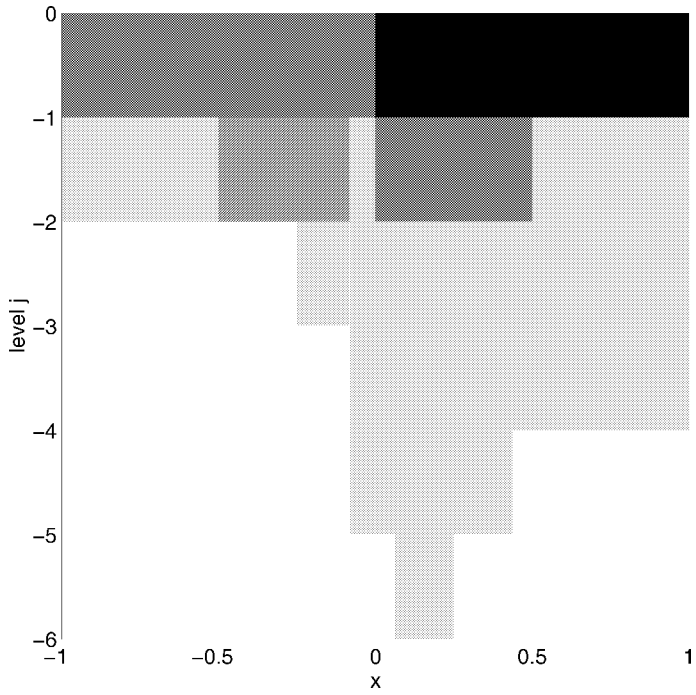

Figure 8. Coefficient distribution for $\mathcal{F}(x)$ with singularity $x_0 = 1/7$ (1-D): The darker the color, the greater the absolute value of the coefficient. $i = 1$, $M = 1$.

of order greater than 1 are not bounded. We can see from figures 2–5 that although the decay is nonmonotonic, it is never so far away from its general tendency at any level that could lead to a false stop.

It can be observed from figures 2–5 that the decay curves for $x_0 = 1/8$ and $x_0 = 1/16$ join the curve for $x_0 = 0$ after several levels. This is because any dyadic point will eventually become a grid point when the grid is refined.

It is also observed from our numerical experiment that the decay is nonmonotonic only in the grid cell containing the singularity.

Figure 7 shows the wavelet coefficient distribution in $[-16, 16]$. This is a pseudo-color plot generated by Matlab. The darker the color, the greater the absolute value. The plot shows in figure 7 is for $i = 1$, $x_0 = 0$. It shows that the higher order wavelet (in this case, $M = 4$) has a faster decay across scales than the lower order one (in this case, $M = 1$).

To see more clearly the change of the wavelet coefficients in the support containing a singularity, we shrink the domain to $[-1, 1]$. Figure 8 shows the coefficient distribution when $x_0 = 1/7$. It can be seen that the coefficients decay rapidly away from the singularity and slower around the singularity. Again it can be observed that the decay in the grid cell containing the singularity is not monotone, e.g., the coefficient is smaller on level 2 than on level 3.

All these observations coincide with our theoretical analysis.

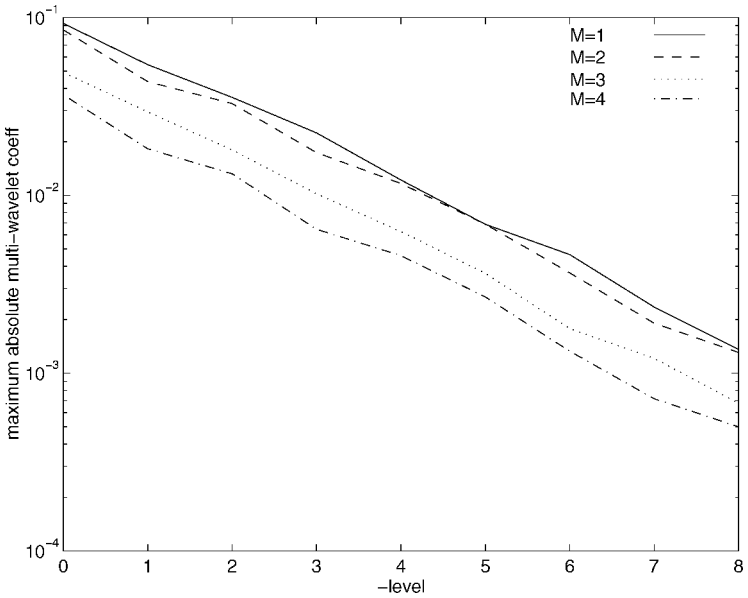

Figure 9. Upper bound envelopes of the decay curves for different M.

Table 2
The upper bound line $\log_2(\max_{i,k} |d_i^{j,k}|) = kj + \log_2 c$ for a fixed M.

Order M	Slope k	Intercept c
1	0.76	0.11
2	0.75	0.09
3	0.80	0.05
4	0.83	0.04
Theory	0.833	

It can be seen from figures 2–5 that the decay curves for all possible x_0 are uniformly bounded for a fixed M. Figure 9 shows such a upper bound envelope for $M = 1, 2, 3$ and 4, respectively. It was generated by plotting the maximum value of curves with singularities at a number of different locations. Furthermore, we would like to give a quantitative measurement as how fast the decay is. This can be represented by the slope k such that

$$\log_2\Big(\max_{i,k} \big|d_i^{j,k}\big|\Big) \leqslant kj + \log_2 c,$$

where c is the intercept. Table 2 shows the numerical results. Theoretically, since

$$\max_{i,k} \big|d_i^{j,k}\big| \leqslant C\delta_j^{5/6},$$

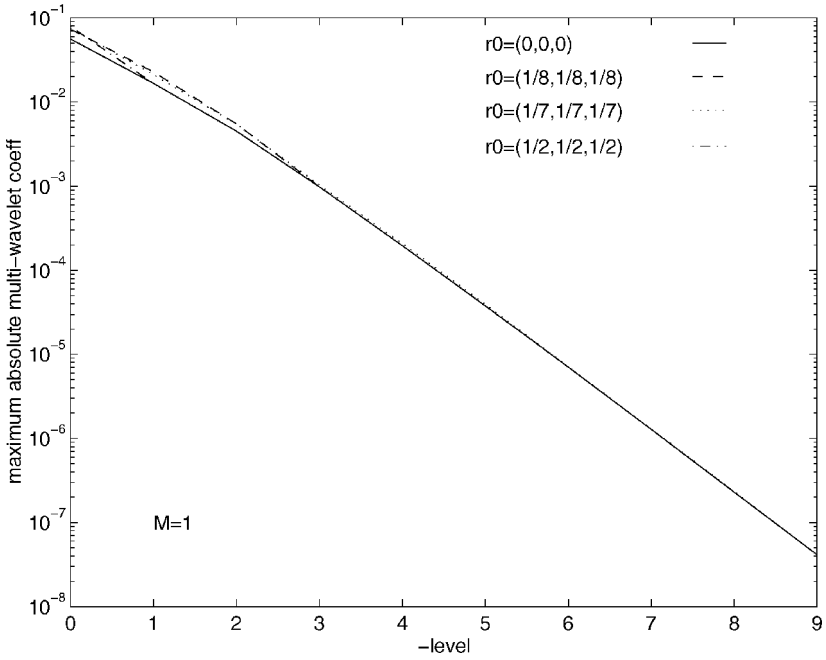

Figure 10. Decay across scales for $M = 1$ (3-D).

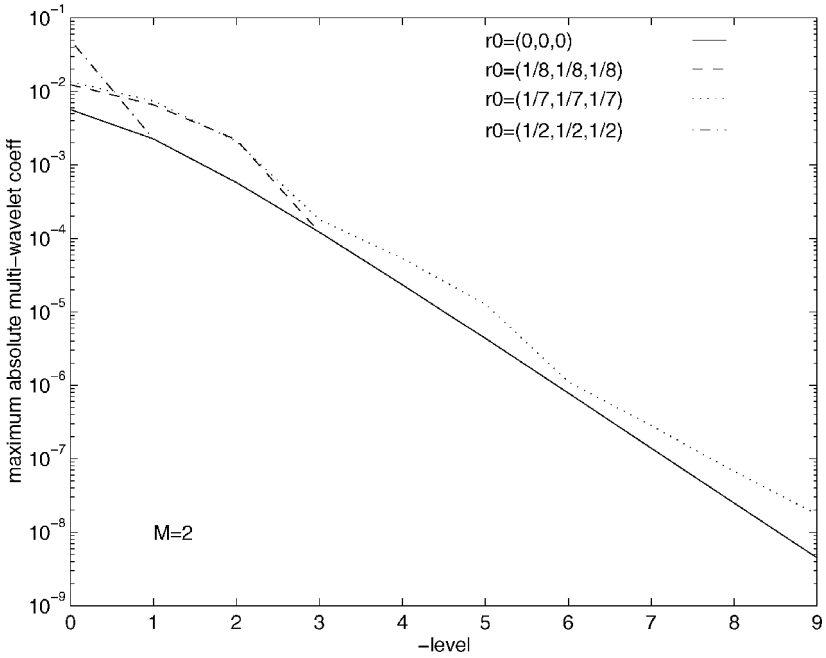

Figure 11. Decay across scales for $M = 2$ (3-D).

where C is some constant, we know that

$$\log_2\Big(\max_{i,k}\big|d_i^{j,k}\big|\Big) \leqslant \frac{5}{6}j + \log_2 C \quad \text{as } j \to -\infty.$$

So the intercept value c in table 2 gives us an idea about the order constant C. Notice that in the table c decreases as M increases.

5.2. *The three-dimensional case*

In this section, we test the 3-D function $\mathcal{F}(r) = \exp(-|r|)$, where

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2},$$

and compare it with the 1-D case.

Figures 10 and 11 show ($\max_{\mathbf{i},\mathbf{k}} |d_{\mathbf{i}}^{j,\mathbf{k}}|$) v.s. level ($-j$) for $M = 1$ and 2, respectively. A number of points (x_0, y_0, z_0)s are tested. Comparing them with the plots in figures 2–5, we see that the 3-D plots decay faster. In theory we know it should be three times as fast as in 1-D. And as in 1-D, the curves with dyadic singularities concur with the curve with the singularity at origin after several levels. We can still observe the oscillations of the curve when the singularity does not concur with grid points. When $M = 1$ these oscillations are not severe due to the multi-dimension cancellation.

Then figure 12 shows the wavelet coefficient distribution in $[-16, 16]$ for $(x_0, y_0, z_0) = (0, 0, 0)$, $y \equiv 0$ and $z \equiv 0$. And figure 13 shows the the coefficient

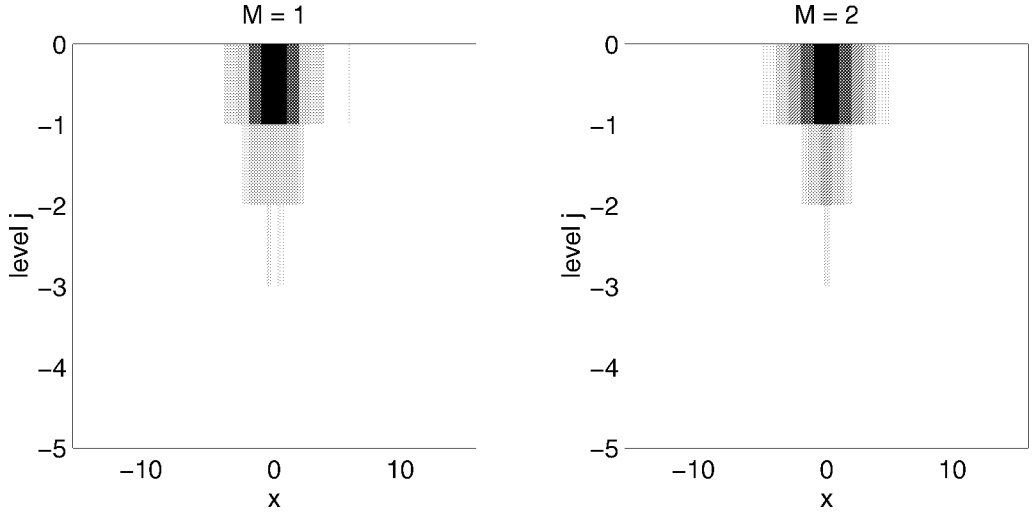

Figure 12. Wavelet coefficient distribution (3-D): In each plot, the darker the color, the greater the absolute value of the coefficient. But for the same darkness, the value in the plot when $M = 1$ is about 10 times the value in the plot when $M = 2$. $\mathbf{i} = (2, 1, 1)$ for $M = 1$, $\mathbf{i} = (3, 1, 1)$ for $M = 2$, $(x_0, y_0, z_0) = (0, 0, 0)$.

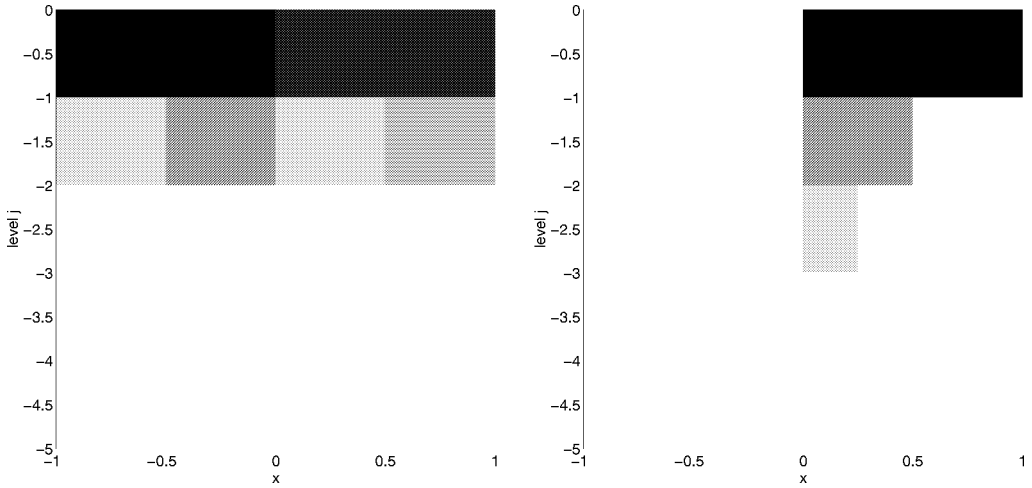

Figure 13. Wavelet coefficient distribution for $\mathcal{F}(x)$ with singularity $(x_0, y_0, z_0) = (1/7, 1/7, 1/7)$ (3-D): The darker the color, the greater the absolute value of the coefficient. $M = 1$ for the plot on the left, $M = 2$ for the plot on the right. $\mathbf{i} = (2, 1, 1)$ for $M = 1$, $\mathbf{i} = (3, 1, 1)$ for $M = 2$.

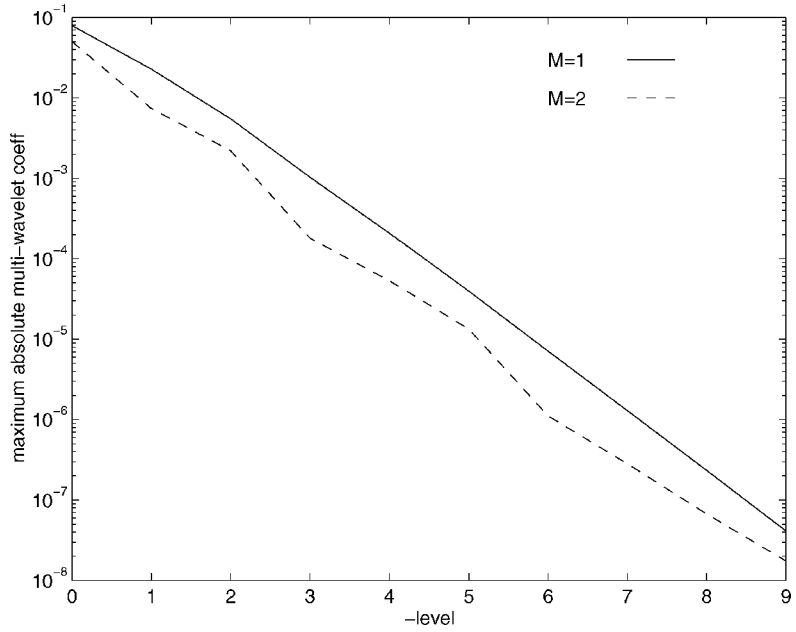

Figure 14. Upper bound envelopes of the decay curves for different M (3-D).

distribution when $(x_0, y_0, z_0) = (1/7, 1/7, 1/7)$, $y \equiv 1/7$ and $z \equiv 1/7$. As in 1-D, the coefficients decay rapidly away from the singularity and slower around the singularity, higher order wavelet has a faster decay across scales than the lower order one. Notice that the maximum value is always obtained near the singularity but does not have to

Table 3
The upper bound line $\log_2(\max_{\mathbf{i},\mathbf{k}} |d_{\mathbf{i}}^{j,\mathbf{k}}|) = kj + \log_2 c$ for a fixed M.

Order M	Slope k	Intercept c
1	2.49	0.24
2	2.42	0.06
Theory	2.50	

be in the interval containing the singularity. This is typical when $M = 1$ due to the cancellation introduced by multiple dimensions.

Finally, figure 14 shows the upper bound envelope for $M = 1$ and 2. Table 3 shows the slope k such that

$$\log_2\left(\max_{\mathbf{i},\mathbf{k}} \left|d_{\mathbf{i}}^{j,\mathbf{k}}\right|\right) \leqslant kj + \log_2 c.$$

We see that the slope in table 3 is about three times the slope in table 2 as we expect. Also, the intercept c decreases significantly from $m = 1$ to $m = 2$, indicating the order constant decays when the order M increases, as we observed in the 1-D case.

6. Conclusion

We have presented a 1-D benchmark singular integral atomic SCF problem which will be useful in the development of multiresolution algorithms for the actual 3-D problems of electronic structure calculations. The 1-D prototype captures the main characteristics of the 3-D problems both theoretically and numerically. The multiwavelet multi-resolution analysis (MRA) is suggested to solve the problem. Theoretical estimates of decay across scales and spatial distribution of wavelet coefficients are given and are confirmed by numerical results in both 1-D and 3-D.

Work is in progress to numerically solve the 1-D prototype problem using fast wavelet algorithms. The application of the operators appearing in the equation can be performed by well-established fast algorithms [BCR], especially the non- standard form. The applicability of these methods to the solution of linear operator problems has been established. However, the equation to be solved here is a linear (1-electron case) or nonlinear (2-electron case) eigenvalue problem. The determination of eigenvalues through multiresolution methods has been recently investigated [2]. We are investigating the application of these methods to the 1-D prototype problem we have presented here.

Acknowledgements

We wish to thank Robert Harrison of the Pacific Northwest National Laboratory for many helpful discussions.

References

[1] B.K. Alpert, Sparse representation of smooth linear operators, Ph.D. thesis, Yale University (1990).
[2] G. Beylkin and N. Coult, A multiresolution stratey for reduction of elliptic PDE's and eigenvalue problems, preprint.
[3] P. Fischer and M. DeFranceschi, Representation of the atomic Hartree–Fock equations in a wavelet basis by means of the BCR algorithm, in: *Wavelets: Theory, Algorithms, and Applications*, eds. C.K. Chui, L. Montefusco and L. Puccio (Academic Press, 1994) pp. 495–506.
[4] P. Fischer and M. DeFranceschi, to appear in SIAM Journal of Numerical Analysis.
[5] Jaffard, Exposants de Hölder en des Points donnés et coéfficients d'ondelettes, C. R. Acad. Sci. Paris, Série 1 308 (1989) 79–81.
[6] M. Karplus and R.N. Porter, *Atoms and Molecules: An Introduction for Students in Physical Chemistry* (W.A. Benjamin, Menlo Park, CA, 1970).
[7] S. Mallat, Multiresolution approximation and wavelets, Trans. Amer. Math. Soc. 315 (1989) 69–88.
[8] Y. Meyer, Ondelettes sur l'intervalle, Revista Matematica Iberoamericana 7 (1991) 115–133.
[9] P.J. Taylor, San Diego Supercomputer Center, private communication (1995).