

## BEST $N$ -TERM APPROXIMATION IN ELECTRONIC STRUCTURE CALCULATIONS

### I. ONE-ELECTRON REDUCED DENSITY MATRIX

HEINZ-JÜRGEN FLAD<sup>1</sup>, WOLFGANG HACKBUSCH<sup>1</sup> AND REINHOLD SCHNEIDER<sup>2</sup>

**Abstract.** We discuss best  $N$ -term approximation spaces for one-electron wavefunctions  $\phi_i$  and reduced density matrices  $\rho$  emerging from Hartree-Fock and density functional theory. The approximation spaces  $A_q^\alpha(H^1)$  for anisotropic wavelet tensor product bases have been recently characterized by Nitsche in terms of tensor product Besov spaces. We have used the norm equivalence of these spaces to weighted  $\ell_q$  spaces of wavelet coefficients to prove that both  $\phi_i$  and  $\rho$  are in  $A_q^\alpha(H^1)$  for all  $\alpha > 0$  with  $\alpha = \frac{1}{q} - \frac{1}{2}$ . Our proof is based on the assumption that the  $\phi_i$  possess an asymptotic smoothness property at the electron-nuclear cusps.

**Mathematics Subject Classification.** 41A50, 41A63, 65Z05, 81V70.

Received: May 11, 2005. Revised: August 11, 2005.

#### 1. INTRODUCTION

The nonrelativistic Schrödinger equation within the Born-Oppenheimer approximation provides a firm basis for most applications in quantum chemistry. For a time independent problem, the stationary solutions correspond to the eigenvalue problem

$$H\Psi_i = E_i\Psi_i, \text{ with } \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \in H^1(\mathbb{R}^{3N}). \quad (1.1)$$

The Hamiltonian in atomic units

$$H = -\frac{1}{2} \sum_{i=1}^N \Delta_i - \sum_{a=1}^K \sum_{i=1}^N \frac{Z_a}{|\mathbf{x}_i - \mathbf{R}_a|} + \sum_{i < j} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} + \sum_{a < b} \frac{Z_a Z_b}{|\mathbf{R}_a - \mathbf{R}_b|},$$

includes Coulomb interactions between electrons and nuclei, where  $\mathbf{x}_i \in \mathbb{R}^3$  are the electron coordinates and  $Z_a$ ,  $\mathbf{R}_a$  are charges and positions of the nuclei, respectively. Due to the presence of singular Coulomb potentials in the Hamiltonian, certain types of inter-particle cusps have to be considered for exact or approximate solutions of the Schrödinger equation (1.1). These cusps correspond to coalescence points of electrons and nuclei. The

---

*Keywords and phrases.* Best  $N$ -term approximation, wavelets, Hartree-Fock method, density functional theory.

<sup>1</sup> Max-Planck-Institut für Mathematik in den Naturwissenschaften, Inselstr. 22-26, 04103 Leipzig, Germany.  
flad@mis.mpg.de; wh@mis.mpg.de

<sup>2</sup> Christian-Albrechts-Universität Kiel, Christian-Albrechts-Platz 4, 24098 Kiel, Germany. rs@numerik.uni-kiel.de

asymptotic behaviour of wavefunctions near cusps has been studied in references [17, 18, 20]. From a computational point of view, the regularity properties of wavefunctions at cusps determine the approximation error with respect to a given basis set. Despite their computational significance rather few rigorous results exist concerning approximation errors for certain types of basis functions. Even for the most popular *Gaussian type orbitals* (GTO) basis functions rigorous results for electron-nuclear and electron-electron cusps exist only in the case of atoms, see references [1, 21] and [16, 22], respectively. Recently, hierarchical tensor product bases, so called sparse grids [2] and hyperbolic wavelets [6] have emerged as an alternative to GTO basis sets in electronic structure calculations [7, 8, 12, 24]. Interesting results for tensor product approximations of exact solutions of the Schrödinger equation (1.1) have been recently obtained by Yserentant [29, 30]. We pursue a more restricted approach by studying approximation errors of hyperbolic wavelet bases for certain classes of approximate wavefunctions. Within the present work, we focus on effective independent-particle models like *density functional theory* (DFT) and the *Hartree-Fock* (HF) method. The latter provides a starting point for many-particle methods that enable the description of electron correlations [15]. Because of the absence of many-electron cusps in these models, we can restrict ourselves to the electron-nuclear cusp, where a single electron approaches a nucleus. Electron-nuclear cusps are the only cusps that have a direct counterpart in the electron density and therefore some significance in the framework of DFT. This has been recently demonstrated in a series of papers [10, 11, 19], where it was proven that the exact electron density is smooth and even analytic away from the nuclei.

### 1.1. Brief outline of best $N$ -term approximation for tensor product bases

The concept of best  $N$ -term approximation belongs to the realm of nonlinear approximation theory. For a detailed exposition of this subject we refer to reference [4]. Its basic idea is to characterize function spaces which possess certain approximation properties with respect to a given basis  $\{\zeta_i : i \in \Lambda\}$  in a separable Hilbert space  $H$ . To be more explicit, we define nonlinear submanifolds

$$\Sigma_N := \left\{ \sum_{i \in \Delta} c_i \zeta_i : \Delta \subset \Lambda, \#\Delta \leq N \right\},$$

which consists of all possible linear combinations of basis functions with at most  $N$  terms. The corresponding approximation error for an arbitrary function  $f \in H$

$$\sigma_N(f) := \inf_{f_N \in \Sigma_N} \|f - f_N\|_H,$$

characterizes the approximation spaces of best  $N$ -term approximations

$$A_q^\alpha(H) := \{f \in H : |f|_{A_q^\alpha(H)} < \infty\}, \quad (1.2)$$

with

$$|f|_{A_q^\alpha(H)} := \left( \sum_{N \in \mathbb{N}} (N^\alpha \sigma_N(f))^q N^{-1} \right)^{\frac{1}{q}}.$$

Loosely speaking, if  $f \in A_q^\alpha(H)$  it is possible to achieve a convergence rate  $\sigma_N(f) \sim N^{-\alpha}$  with respect to the number of basis functions. Here and in the following  $a \sim b$  means that the quantities can be uniformly bounded by some constant multiple of each other. Similarly  $a \lesssim b$  means that  $a$  is uniformly bounded by some constant multiple of  $b$ . For our envisaged applications in electronic structure calculations, we consider  $L^2(\Omega)$  and the Sobolev space  $H^1(\Omega)$  on bounded domains  $\Omega \subset \mathbb{R}^d$  equipped with their natural Hilbert space structure.

It is not our intention to give a systematic exposition of best  $N$ -term approximations. Instead we restrict ourselves to orthogonal wavelet bases and highlight on the role of dimension  $d$ . In the one-dimensional case of a bounded interval  $I \subset \mathbb{R}$ , wavelets  $\psi_{j,a} := 2^{j/2} \psi(2^j x - a)$  with  $j \geq j_0$  and scaling functions

$\varphi_{j_0,a} := 2^{j_0/2}\varphi(2^{j_0}x - a)$  on the coarsest level  $j_0$  provide an orthogonal basis in  $L^2(I)$ . According to DeVore, Jawerth and Popov [5], the approximation spaces (1.2) can be identified with Besov spaces

$$A_q^{\alpha-t}(H) = B_q^\alpha(L_q(I)), \text{ if } \alpha = t + \frac{1}{q} - \frac{1}{2}, \quad (1.3)$$

where  $t = 0, 1$  for  $H = L^2(I)$ ,  $H^1(I)$ , respectively. The Besov spaces are norm equivalent to weighted  $\ell_q$  spaces of wavelet coefficients

$$\|f\|_{B_q^\alpha(L_q(I))}^q \sim \sum_a |\langle \varphi_{j_0,a} | f \rangle|^q + \sum_{j \geq j_0} 2^{jq(\alpha - \frac{1}{q} + \frac{1}{2})} \left( \sum_b |\langle \psi_{j,b} | f \rangle|^q \right), \text{ for } \alpha, q > 0, \quad (1.4)$$

if (i)  $\psi, \varphi \in B_q^\beta(L_q(I))$  for some  $\beta > \alpha$ , (ii)  $\psi$  has  $p$  vanishing moments with  $p > \alpha$ , (iii)  $p > 1/q - 1$  if  $q < 1$ .

There exist two different approaches in order to extend (1.3) to higher dimensions. Within the first approach, multivariate isotropic wavelets

$$\gamma_{\mathbf{j},\mathbf{a}}^{(\mathbf{s})}(\mathbf{x}) = \psi_{j_1,a_1}^{(s_1)}(x_1) \psi_{j_2,a_2}^{(s_2)}(x_2) \dots \psi_{j_d,a_d}^{(s_d)}(x_d), \text{ with } \mathbf{s} := (s_1, s_2, \dots, s_d), \mathbf{a} := (a_1, a_2, \dots, a_d), \quad (1.5)$$

are obtained by taking mixed tensor products of univariate scaling functions  $\psi_{j,a}^{(0)}(x) := \varphi_{j,a}(x)$  and wavelets  $\psi_{j,a}^{(1)}(x) := \psi_{j,a}(x)$  on the same level of refinement  $j$ . Pure scaling function tensor products  $\gamma_{j_0,\mathbf{a}}^{(\mathbf{0})}$  are included on the coarsest level  $j_0$ . For such kind of wavelet bases it has been shown [3] that the approximation spaces (1.2) correspond to Besov spaces

$$A_q^{(\alpha-t)/d}(H) = B_q^\alpha(L_q(\Omega)), \text{ if } \alpha = t + \frac{d}{q} - \frac{d}{2}, \quad (1.6)$$

where  $t = 0, 1$  for  $H = L^2(\Omega)$ ,  $H^1(\Omega)$ , respectively. These Besov spaces are again norm equivalent to weighted  $\ell_q$  spaces of wavelet coefficients

$$\|f\|_{B_q^\alpha(L_q(\Omega))}^q \sim \sum_{j \geq j_0} \sum_{\mathbf{s}} 2^{jq(\alpha - \frac{d}{q} + \frac{d}{2})} \left( \sum_{\mathbf{a}} |\langle \gamma_{j,\mathbf{a}}^{(\mathbf{s})} | f \rangle|^q \right) \text{ for } \alpha, q > 0, \quad (1.7)$$

provided that the underlying univariate wavelet basis satisfies requirements (i) to (iii).

Due to the dimensional dependence of equation (1.6), isotropic wavelet tensor products loose much of their efficiency in higher dimensions. At present, the only practicable approach to high dimensional problems are hyperbolic wavelets [6] where the multivariate anisotropic wavelet basis

$$\chi_{\mathbf{j},\mathbf{a}}(\mathbf{x}) = \psi_{j_1,a_1}^{(s_1)}(x_1) \psi_{j_2,a_2}^{(s_2)}(x_2) \dots \psi_{j_d,a_d}^{(s_d)}(x_d), \text{ with } \mathbf{j} := ((j_1, s_1), (j_2, s_2), \dots, (j_d, s_d)), \quad (1.8)$$

consists of tensor products of wavelets on possibly different levels. Univariate scaling functions  $\psi_{j_0,a}^{(0)}$  may appear in these tensor products on the coarsest level  $j_0$  only. Such kind of hierarchical tensor product bases are well known in finite element methods as sparse grids [2]. Recently, best  $N$ -term approximation has been extended to anisotropic wavelet bases by Nitsche [25]. The corresponding approximation spaces (1.2) were characterized by Nitsche in terms of suitably defined tensor product Besov spaces. We restrict our discussion to the Hilbert space  $H^1(\Omega)$  on a cube  $\Omega = I_1 \times I_2 \times \dots \times I_d \subset \mathbb{R}^d$  and refer for  $L^2(\Omega)$  to Nitsche's paper. In this case it is necessary to consider tensor product Besov spaces

$$\mathcal{B}_{q,i}^\alpha(\Omega) = \bigotimes_{k=1}^d B_q^{\alpha+\delta_{i,k}}(L_q(I_k)),$$

which have a distinguished direction. These spaces are norm equivalent to weighted  $\ell_q$  norms for anisotropic wavelet coefficients

$$\|f\|_{\mathcal{B}_{q,i}^\alpha}^q \sim \sum_{\mathbf{j}} 2^{j_i q} \left( \sum_{\mathbf{a}} |\langle \chi_{\mathbf{j},\mathbf{a}} | f \rangle|^q \right), \text{ if } \frac{1}{q} = \alpha + \frac{1}{2}.$$

The appropriate Besov spaces for best  $N$ -term approximation can be obtained as intersections

$$\mathcal{B}_q^\alpha(\Omega) = \bigcap_i \mathcal{B}_{q,i}^\alpha(\Omega).$$

A function  $f$  belongs to  $\mathcal{B}_q^\alpha$  if and only if it is bounded with respect to the norm

$$\|f\|_{\mathcal{B}_q^\alpha}^q = \sum_{\mathbf{j}} 2^{\max\{j_i\}q} \left( \sum_{\mathbf{a}} |\langle \chi_{\mathbf{j},\mathbf{a}} | f \rangle|^q \right). \quad (1.9)$$

This norm requires univariate wavelets with  $p > \alpha + 1$  vanishing moments. The corresponding relation between approximation and Besov spaces is given by

$$A_q^\alpha(H^1(\Omega)) = \mathcal{B}_q^\alpha(\Omega), \text{ if } \alpha = \frac{1}{q} - \frac{1}{2}.$$

## 2. BEST $N$ -TERM APPROXIMATION FOR HF AND DFT METHODS

In the present work, we want to study best  $N$ -term approximation for one-electron wavefunctions  $\phi_i(\mathbf{x})$ , so called orbitals, and one-electron reduced density matrices  $\rho(\mathbf{x}, \mathbf{y})$  which appear in HF or DFT methods. The latter can be expressed in terms of orbitals

$$\rho(\mathbf{x}, \mathbf{y}) = \sum_i n_i \phi_i(\mathbf{x}) \phi_i^*(\mathbf{y}),$$

where  $n_i$  is the occupation number of the  $i$ th orbital. In HF and DFT methods only a finite number of  $n_i$  are nonzero. Therefore, we can focus our discussion on orbitals, which in the case of HF orbitals are known to be  $C^\infty(\mathbb{R}^3 \setminus A)$  [23] where  $A \subset \mathbb{R}^3$  corresponds to a discrete subset of the positions of the nuclei.

We assume that in an appropriately chosen bounded neighbourhood  $\Omega$  of a nucleus located at  $\mathbf{R}$ , the orbitals possess an asymptotic smoothness property

$$|\partial^\beta \phi_i(\mathbf{x})| \leq c_\beta |\mathbf{x} - \mathbf{R}|^{1-|\beta|}, \text{ for all } \mathbf{x} \neq \mathbf{R} \text{ and all } |\beta| \geq 1,$$

where we have introduced the usual short-hand notation

$$\partial^\beta := \frac{\partial^{\beta_1}}{\partial x_1^{\beta_1}} \frac{\partial^{\beta_2}}{\partial x_2^{\beta_2}} \frac{\partial^{\beta_3}}{\partial x_3^{\beta_3}},$$

with absolute value of the multi-index  $|\beta| := \beta_1 + \beta_2 + \beta_3$ . This is actually a rather weak requirement that is satisfied *e.g.* by the eigenfunctions of the hydrogen atom or by Slater type orbitals (STO) [15]

$$\phi_{nlm}^{\text{STO}}(r, \varphi, \theta) = C_{nlm} Y_{lm}(\varphi, \theta) r^{n-1} e^{-\zeta r}, \text{ for } n \geq 1 \quad (2.1)$$

which are supposed to be the most efficient basis sets for atomic HF calculations. The STOs are expressed in spherical coordinates with  $r = |\mathbf{x} - \mathbf{R}|$ , where the angular dependent part is represented by spherical harmonics  $Y_{lm}$ .

## 2.1. Besov regularity of orbitals near electron-nuclear cusps

Since electron-nuclear cusps are well separated it is sufficient to consider a single nucleus. For simplicity, we assume the nucleus to be located at the origin. In order to determine best  $N$ -term approximation spaces for orbitals we prove the following lemma.

**Lemma 1.** *Suppose  $\phi \in C^\infty(\mathbb{R}^3 \setminus \mathbf{0})$  is asymptotically smooth*

$$|\partial^\beta \phi(\mathbf{x})| \leq c_\beta |\mathbf{x}|^{1-|\beta|}, \text{ for all } \mathbf{x} \neq \mathbf{0} \text{ and all } |\beta| \geq 1 \quad (2.2)$$

in a bounded neighbourhood  $\Omega \subset \mathbb{R}^3$  of the origin. Then  $\phi \in \mathcal{B}_q^\alpha(\Omega)$  for all  $q > 0$  and  $\alpha = \frac{1}{q} - \frac{1}{2}$ .

*Proof.* The prove of this lemma requires some general estimates for wavelet coefficients in various dimensions. For convenience of the reader, we present these estimates with a sketch of their proof. Further details can be found e.g. in [26].

**Proposition 1.** *If  $f \in Lip(\mathbb{R})$ , then*

$$\left| \int_{\mathbb{R}} f(x) \psi_{j,k}(x) dx \right| \lesssim 2^{-\frac{3}{2}j} \|f'\|_{L^\infty(\text{supp } \psi_{j,k})}.$$

*Proof.* Using the identity

$$f(x) = f(x_0) + (x - x_0) \int_0^1 f'(x_0 + t(x - x_0)) dt,$$

together with the vanishing moments property of wavelets we obtain the desired estimate.  $\square$

**Proposition 2.** *Suppose the function  $f(\mathbf{x})$  with  $\mathbf{x} \in \mathbb{R}^d$  is smooth on the support of a pure wavelet tensor product  $\chi_{\mathbf{j},\mathbf{a}}$  i.e.  $\mathbf{j} := ((j_1, 1), (j_2, 1), \dots, (j_d, 1))$ . Then the following estimate holds*

$$\left| \int_{\mathbb{R}^d} f(\mathbf{x}) \chi_{\mathbf{j},\mathbf{a}}(\mathbf{x}) d\mathbf{x} \right| \lesssim 2^{-(p+1/2)|\mathbf{j}|} \|\partial_{x_1}^{j_1} \dots \partial_{x_d}^{j_d} f\|_{L^\infty(\text{supp } \chi_{\mathbf{j},\mathbf{a}})}, \text{ with } |\mathbf{j}| := \sum_{i=1}^d j_i.$$

*Proof.* We proceed in a recursive manner by first considering the case  $d = 1$ . In the integrand, we replace  $f$  by its Taylor series expansion at  $x_0 \in \text{supp } \psi_{j,a}$ . Due to the vanishing moments property, contributions of derivatives  $\partial^n f(x_0)$  with  $n < p$  vanish. The remainder of the Taylor series can be represented by the integral

$$\frac{(x - x_0)^p}{(p-1)!} \int_0^1 (1-t)^{p-1} \partial^p f(x_0 + t(x - x_0)) dt,$$

from which the estimate for  $d = 1$  follows. Through successive application of the  $d = 1$  estimate to each of the variables separately, the corresponding estimates for  $d > 1$  follow.  $\square$

Due to the strongly anisotropic character of the tensor products (1.8) it is necessary to subdivide the parameter set  $\Lambda_{\mathbf{j}} := \{\mathbf{a} \in \mathbb{Z}^3 : (2^{-j_1} a_1, 2^{-j_2} a_2, 2^{-j_3} a_3) \in \Omega\}$ .

**Definition 1.** For each combination of wavelet levels  $j_1 \geq j_2 \geq j_3 \geq j_0$ , the discrete set  $\Lambda_{\mathbf{j}}$  of translation parameters  $\mathbf{a} = (a_1, a_2, a_3)$  is subdivided with respect to the parameter  $\delta := 2L$  into the domains

- (i)  $A_{\mathbf{j}} := \{\mathbf{a} \in \Lambda_{\mathbf{j}} : |a_1|, |a_2|, |a_3| \leq \delta\}$ ;
- (ii)  $B_{\mathbf{j}} := \{\mathbf{a} \in \Lambda_{\mathbf{j}} : \delta < |a_1| \leq 2^{j_1-j_2} \delta; |a_2|, |a_3| \leq \delta\}$ ;
- (iii)  $C_{\mathbf{j}} = C_a \cup C_b$ ;  
 $C_a := \{\mathbf{a} \in \Lambda_{\mathbf{j}} : 2^{j_1-j_2} \delta < |a_1| \leq 2^{j_1-j_3} \delta; |a_2|, |a_3| \leq \delta\}$ ;  
 $C_b := \{\mathbf{a} \in \Lambda_{\mathbf{j}} : |a_1| \leq 2^{j_1-j_3} \delta; \delta < |a_2| \leq 2^{j_2-j_3} \delta; |a_3| \leq \delta\}$ ;
- (iv)  $D_{\mathbf{j}} := \Lambda_{\mathbf{j}} \setminus (A_{\mathbf{j}} \cup B_{\mathbf{j}} \cup C_{\mathbf{j}})$ .

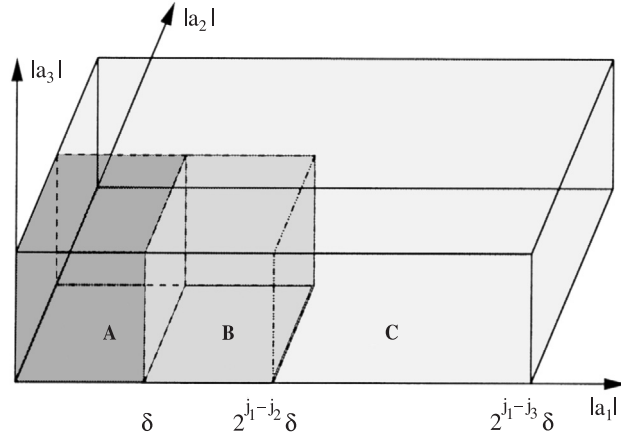


FIGURE 1. Subdivision of the set of translation parameters  $\Lambda_{\mathbf{j}}$ .

The parameter  $L$  is chosen in such a way that  $\text{supp } \psi^{(s)} \subset [-L/2, L/2]$ . See Figure 1 for a schematic graphic representation.

In order to demonstrate the finiteness of the norm (1.9) it is sufficient to consider the case  $j_1 \geq j_2 \geq j_3 \geq j_0$ . With the Definition 1 at hand, we decompose the norm (1.9) into four parts

$$\sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{j_1 q} \left( \sum_{\mathbf{a}} |\langle \chi_{\mathbf{j}, \mathbf{a}} | \phi \rangle|^q \right) = I_A + I_B + I_C + I_D,$$

with

$$I_A = \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{j_1 q} \left( \sum_{\mathbf{a} \in A_{\mathbf{j}}} |\langle \chi_{\mathbf{j}, \mathbf{a}} | \phi \rangle|^q \right), \text{ etc.}$$

where the coarsest level  $j_0$  is chosen in such a way that  $2^{-j_0} \approx \text{diam } \Omega$ .

### Domain A

First we observe that  $\#A_{\mathbf{j}} = O(1)$  with respect to the wavelet levels  $\mathbf{j}$ . Therefore it is sufficient to estimate a single wavelet coefficient  $\langle \chi_{\mathbf{j}, \mathbf{a}} | \phi \rangle$  in this case. In order to apply the asymptotic smoothness property (2.2), we decompose the cuboid  $\Omega_{\mathbf{j}, \mathbf{a}}$  with center at  $\mathbf{b} := (2^{-j_1} a_1, 2^{-j_2} a_2, 2^{-j_3} a_3)$  and edge-lengths  $2^{-j_1} L \times 2^{-j_2} L \times 2^{-j_3} L$  into subcubes  $\square_i$  ( $i \in \Delta$ ) with edge length  $2^{-j_1} L$ . According to Def. (1), we have  $\text{supp } \chi_{\mathbf{j}, \mathbf{a}} \subset \Omega_{\mathbf{j}, \mathbf{a}}$ . The subcubes  $\square_i$  ( $i \in \Delta_0 := \{i \in \Delta, : \text{dist}(\square_i, \mathbf{0}) < 2^{-j_1} L\}$ ) close to the origin are considered separately. Their number  $\#\Delta_0 = O(1)$  is independent of the wavelet levels  $\mathbf{j}$ . For the remaining subcubes  $\square_i$  ( $i \in \Delta \setminus \Delta_0$ ) it becomes necessary to control their contributions with respect to the distance from the origin, because  $\#(\Delta \setminus \Delta_0) = O(2^{j_1 - j_2} 2^{j_1 - j_3})$  depends on the wavelet levels.

In the first step, we decompose the integral

$$\begin{aligned} \left| \int_{\Omega_{\mathbf{j}, \mathbf{a}}} \phi(\mathbf{x}) \chi_{\mathbf{j}, \mathbf{a}}(\mathbf{x}) d\mathbf{x} \right| &= \left| \sum_{i \in \Delta_0} \int_{\square_i} \phi(\mathbf{x}) \chi_{\mathbf{j}, \mathbf{a}}(\mathbf{x}) d\mathbf{x} + \sum_{i \in \Delta \setminus \Delta_0} \int_{\square_i} \phi(\mathbf{x}) \chi_{\mathbf{j}, \mathbf{a}}(\mathbf{x}) d\mathbf{x} \right| \\ &\leq \sum_{i \in \Delta_0} \left| \int_{\square_i} \phi(\mathbf{x}) \chi_{\mathbf{j}, \mathbf{a}}(\mathbf{x}) d\mathbf{x} \right| + \sum_{i \in \Delta \setminus \Delta_0} \left| \int_{\square_i} \phi(\mathbf{b} + \mathbf{c}_i + \mathbf{y}) \chi_{\mathbf{j}, \mathbf{0}}(\mathbf{c}_i + \mathbf{y}) d\mathbf{y} \right|, \quad (2.3) \end{aligned}$$

where we have introduced local coordinates  $\mathbf{y}$  for those subcubes, which are subject to the asymptotic smoothness condition (2.2). Local coordinates  $\mathbf{y}$  are defined *via*  $\mathbf{x} = \mathbf{b} + \mathbf{c}_i + \mathbf{y}$ , where  $\mathbf{b}$  denotes the center of the cuboid  $\Omega_{\mathbf{j}, \mathbf{a}}$  and  $\mathbf{c}_i = 2^{-(j_1+1)}L(0, e_2^{(i)}, e_3^{(i)})$ , with  $e_2^{(i)}, e_3^{(i)} \in \mathbb{Z}$ , points from the center of the cuboid to the center of the subcube  $\square_i$ . This construction is shown schematically in Figure 2.

We can now apply Proposition 2 for the case  $d = 1$  and the asymptotic smoothness condition (2.2) to subcubes  $\square_i$  ( $i \in \Delta \setminus \Delta_0$ )

$$\begin{aligned} \left| \int_{\square_i} \phi(\mathbf{b} + \mathbf{c}_i + \mathbf{y}) \chi_{\mathbf{j}, \mathbf{0}}(\mathbf{c}_i + \mathbf{y}) d\mathbf{y} \right| &\lesssim \underbrace{2^{-(p+\frac{1}{2})j_1}}_{(a)} \underbrace{2^{\frac{1}{2}(j_2+j_3)}}_{(b)} \underbrace{2^{-2j_1}}_{(c)} \|\partial_{x_1}^p \phi\|_{L_\infty(\square_i)} \\ &= 2^{-(p+\frac{5}{2})j_1} 2^{\frac{1}{2}(j_2+j_3)} \|\partial_{x_1}^p \phi\|_{L_\infty(\square_i)} \\ &\lesssim 2^{-(p+\frac{5}{2})j_1} 2^{\frac{1}{2}(j_2+j_3)} |\mathbf{b} + \mathbf{c}_i|^{1-p}, \end{aligned} \quad (2.4)$$

where the prefactor in the first line constitutes in the following way: (a) Proposition 2 for variable  $y_1$ , (b) normalization constant of the wavelet component for variables  $y_2, y_3$ , (c) edge lengths of the subcube  $\square_i$  in  $y_2$  and  $y_3$  direction. The sum can be estimated by

$$\begin{aligned} \sum_{i \in \Delta \setminus \Delta_0} |\mathbf{b} + \mathbf{c}_i|^{1-p} &\lesssim 2^{2j_1} \int_{2^{-j_1}L}^{2^{-j_3}L} r^{2-p} dr \\ &\lesssim \begin{cases} 2^{2j_1} 2^{-j_3}, & \text{if } p = 2 \\ (j_1 - j_3 + 1) 2^{2j_1}, & \text{if } p = 3 \\ 2^{(p-1)j_1}, & \text{if } p > 3. \end{cases} \end{aligned} \quad (2.5)$$

Combining both estimates (2.4) and (2.5), we get bounds for the second sum in the decomposition (2.3)

$$\sum_{i \in \Delta \setminus \Delta_0} \left| \int_{\square_i} \phi(\mathbf{x}) \chi_{\mathbf{j}, \mathbf{a}}(\mathbf{x}) d\mathbf{x} \right| \lesssim \begin{cases} 2^{-\frac{5}{2}j_1} 2^{\frac{1}{2}j_2} 2^{-\frac{1}{2}j_3}, & \text{if } p = 2 \\ (j_1 - j_3 + 1) 2^{-\frac{7}{2}j_1} 2^{\frac{1}{2}(j_2+j_3)}, & \text{if } p = 3 \\ 2^{-\frac{7}{2}j_1} 2^{\frac{1}{2}(j_2+j_3)}, & \text{if } p > 3. \end{cases}$$

It remains to estimate the first sum in the decomposition (2.3)

$$\begin{aligned} \sum_{i \in \Delta_0} \left| \int_{\square_i} \phi(\mathbf{x}) \chi_{\mathbf{j}, \mathbf{a}}(\mathbf{x}) d\mathbf{x} \right| &\lesssim \sum_{i \in \Delta_0} 2^{-\frac{7}{2}j_1} 2^{\frac{1}{2}(j_2+j_3)} \|\partial_{x_1}^p \phi\|_{L_\infty(\square_i)} \\ &\lesssim 2^{-\frac{7}{2}j_1} 2^{\frac{1}{2}(j_2+j_3)}, \end{aligned}$$

where we have used  $\#\Delta_0 = O(1)$ , Proposition 1 and the Lipschitz continuity of  $\phi$ .

Putting things together, we obtain for domain A the final estimate

$$\begin{aligned} I_A &= \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{j_1 q} \left( \sum_{\mathbf{a} \in A_{\mathbf{j}}} |\langle \chi_{\mathbf{j}, \mathbf{a}} | \phi \rangle|^q \right) \\ &\lesssim \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} \begin{cases} 2^{-\frac{3}{2}qj_1} 2^{\frac{1}{2}qj_2} 2^{-\frac{1}{2}qj_3}, & \text{if } p = 2 \\ (j_1 - j_3 + 1)^q 2^{-\frac{5}{2}qj_1} 2^{\frac{1}{2}q(j_2+j_3)}, & \text{if } p = 3 \\ 2^{-\frac{5}{2}qj_1} 2^{\frac{1}{2}q(j_2+j_3)}, & \text{if } p > 3 \end{cases} \\ &< \infty \end{aligned}$$

which proves the finiteness of the first part of the norm (1.9).

### Domain B

This case can be treated along the same line as the previous one. The cuboids  $\Omega_{\mathbf{j}, \mathbf{a}}$  are again decomposed into subcubes with edge length  $2^{-j_1} L$ . According to Definition 1, each subcube satisfies  $\text{dist}(\square_i, \mathbf{0}) > 2^{-j_1} L$ . For a single wavelet coefficient, we obtain the estimate

$$\begin{aligned} \left| \int_{\Omega_{\mathbf{j}, \mathbf{a}}} \phi(\mathbf{x}) \chi_{\mathbf{j}, \mathbf{a}}(\mathbf{x}) d\mathbf{x} \right| &\leq \sum_{i \in \Delta} \left| \int_{\square_i} \phi(\mathbf{x}) \chi_{\mathbf{j}, \mathbf{a}}(\mathbf{x}) d\mathbf{x} \right| \\ &\lesssim 2^{-(p+\frac{5}{2})j_1} 2^{\frac{1}{2}(j_2+j_3)} \sum_{i \in \Delta} |\mathbf{b} + \mathbf{c}_i|^{1-p} \\ &\lesssim 2^{-(p+\frac{1}{2})j_1} 2^{-\frac{1}{2}(j_2+j_3)} |b_1|^{1-p}, \end{aligned} \quad (2.6)$$

from the previous estimate (2.4), where  $\#\Delta = O(2^{j_1-j_2} 2^{j_1-j_3})$  has been used in the third line. Summing up the whole domain, we get the bounds

$$\begin{aligned} \sum_{\mathbf{a} \in B_{\mathbf{j}}} |b_1|^{-q(p-1)} &\lesssim 2^{q(p-1)j_1} \sum_{\delta \leq a_1 \leq 2^{j_1-j_2}\delta} |a_1|^{-q(p-1)} \\ &\lesssim 2^{q(p-1)j_1} \int_{\delta}^{2^{j_1-j_2}\delta} |a_1|^{-q(p-1)} da_1 \\ &\lesssim \begin{cases} 2^{j_1} 2^{-(1-q(p-1))j_2}, & \text{if } 1 - \frac{q}{2} < q(p-1) < 1 \\ (j_1 - j_2 + 1) 2^{j_1}, & \text{if } q(p-1) = 1 \\ 2^{q(p-1)j_1}, & \text{if } q(p-1) > 1, \end{cases} \end{aligned} \quad (2.7)$$

depending on the parameters  $p, q$ . Combining estimates (2.6) and (2.7), yields the estimate

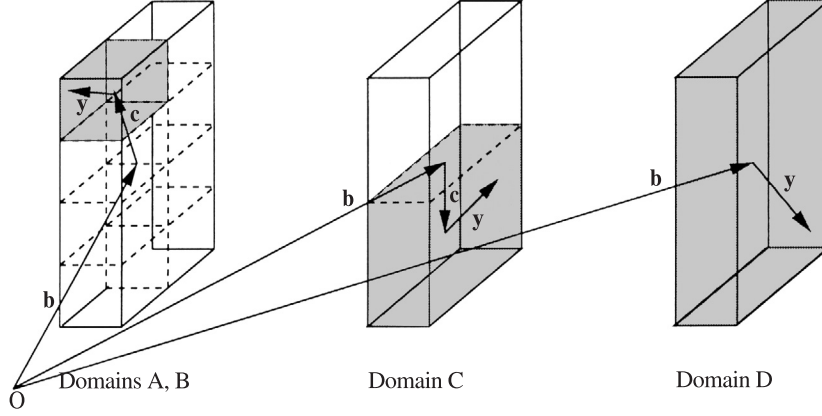
$$\begin{aligned} I_B &= \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{j_1 q} \left( \sum_{\mathbf{a} \in B_{\mathbf{j}}} |\langle \chi_{\mathbf{j}, \mathbf{a}} | \phi \rangle|^q \right) \\ &\lesssim \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} \begin{cases} 2^{-(pq-1-\frac{q}{2})j_1} 2^{-\frac{q}{2}(j_2+j_3)}, & \text{if } 1 - \frac{q}{2} < q(p-1) < 1 \\ (j_1 - j_2 + 1) 2^{-\frac{q}{2}j_1} 2^{-\frac{q}{2}(j_2+j_3)}, & \text{if } q(p-1) = 1 \\ 2^{-\frac{q}{2}j_1} 2^{-\frac{q}{2}(j_2+j_3)}, & \text{if } q(p-1) > 1 \end{cases} \\ &< \infty \end{aligned}$$

which demonstrates the finiteness of the contribution from wavelet coefficients in domain B to the norm (1.9).

### Domain C

Cuboids  $\Omega_{\mathbf{j}, \mathbf{a}}$  that belong to domain C are decomposed into subcuboids  $\square_i$  ( $i \in \Delta$ ) with edge lengths  $2^{-j_1} L \times 2^{-j_2} L \times 2^{-j_3} L$ . The decomposition is shown schematically in Figure 2. According to Definition 1, each subcube satisfies  $\text{dist}(\square_i, \mathbf{0}) > 2^{-j_2} L$ . The proof follows essentially the same line of arguments as in the previous cases. We can benefit, however, in this case from vanishing moments in two directions.



FIGURE 2. Domain decomposition of cuboids  $\Omega_{j,\mathbf{a}}$ .

Using Proposition 2 with  $d = 2$ , we can estimate the wavelet coefficients

$$\begin{aligned}
\left| \int_{\Omega_{j,\mathbf{a}}} \phi(\mathbf{x}) \chi_{j,\mathbf{a}}(\mathbf{x}) d\mathbf{x} \right| &= \left| \sum_{i \in \Delta} \int_{\square_i} \phi(\mathbf{b} + \mathbf{c}_i + \mathbf{y}) \chi_{j,\mathbf{0}}(\mathbf{c}_i + \mathbf{y}) d\mathbf{y} \right| \\
&\lesssim 2^{-(p+\frac{1}{2})j_1} 2^{-(p+\frac{3}{2})j_2} 2^{\frac{1}{2}j_3} \sum_{i \in \Delta} \|\partial_{x_1}^p \partial_{x_2}^p \phi\|_{L_\infty(\square_i)} \\
&\lesssim 2^{-(p+\frac{1}{2})j_1} 2^{-(p+\frac{3}{2})j_2} 2^{\frac{1}{2}j_3} \sum_{i \in \Delta} |\mathbf{b} + \mathbf{c}_i|^{1-2p} \\
&\lesssim 2^{-(p+\frac{1}{2})(j_1+j_2)} 2^{-\frac{1}{2}j_3} (b_1^2 + b_2^2)^{\frac{1}{2}-p}, \tag{2.8}
\end{aligned}$$

where the prefactor in the first line constitutes analogously to estimate (2.4) by applying Proposition 2 to the variables  $y_1, y_2$  and taking an edge length of  $2^{-j_2}L$  for subcuboids  $\square_i$  in the  $y_3$  direction. Furthermore we have used  $\#\Delta = O(2^{j_2-j_3})$  in the fourth line. The sum over domain C can be majorized by an integral

$$\begin{aligned}
\sum_{\mathbf{a} \in C_j} (b_1^2 + b_2^2)^{-q(p-\frac{1}{2})} &= \sum_{\mathbf{a} \in C_j} |2^{-2j_1} a_1^2 + 2^{-2j_2} a_2^2|^{-q(p-\frac{1}{2})} \\
&\lesssim 2^{j_1+j_2} \int_{|\tilde{\mathbf{a}}| \geq 2^{-j_2} \delta} |\tilde{\mathbf{a}}|^{-q(2p-1)} d\tilde{\mathbf{a}} \\
&\lesssim 2^{j_1} 2^{((2p-1)q-1)j_2}, \tag{2.9}
\end{aligned}$$

because of the inequality  $(2p-1)q-1 > 1$ . Putting things together, we obtain from estimates (2.8) and (2.9) an upper bound for the contributions of wavelet coefficients from domain C

$$\begin{aligned}
I_C &= \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{j_1 q} \left( \sum_{\mathbf{a} \in C_j} |\langle \chi_{j,\mathbf{a}} | \phi \rangle|^q \right) \\
&\lesssim \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{-\epsilon j_1} 2^{-(pq-1-\frac{q}{2}-\epsilon)(j_1-j_2)} 2^{-(q-\epsilon)j_2} 2^{-\frac{q}{2}j_3} \\
&\lesssim \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{-\epsilon j_1} 2^{-(q-\epsilon)j_2} 2^{-\frac{q}{2}j_3} \\
&< \infty,
\end{aligned}$$

where an arbitrary parameter  $0 < \epsilon < \min\{pq-1-\frac{q}{2}, q\}$  has been introduced in order to demonstrate convergence in the norm (1.9).

### Domain D

In the last case, the vanishing moments property applies to all three directions. All cuboids  $\Omega_{\mathbf{j}, \mathbf{a}}$  in domain D with edge lengths  $2^{-j_1}L \times 2^{-j_2}L \times 2^{-j_3}L$  satisfy the distance criteria  $\text{dist}(\Omega_{\mathbf{j}, \mathbf{a}}, \mathbf{0}) > 2^{-j_3}L$  according to Definition 1. The wavelet coefficients can be estimated by

$$\begin{aligned} \left| \int_{\Omega_{\mathbf{j}, \mathbf{a}}} \phi(\mathbf{x}) \chi_{\mathbf{j}, \mathbf{a}}(\mathbf{x}) d\mathbf{x} \right| &= \left| \int_{\Omega_{\mathbf{j}, \mathbf{a}}} \phi(\mathbf{b} + \mathbf{y}) \chi_{\mathbf{j}, \mathbf{0}}(\mathbf{y}) d\mathbf{y} \right| \\ &\lesssim 2^{-(p+\frac{1}{2})(j_1+j_2+j_3)} \|\partial_{x_1}^p \partial_{x_2}^p \partial_{x_3}^p \phi\|_{L^\infty(\Omega_{\mathbf{j}, \mathbf{a}})} \\ &\lesssim 2^{-(p+\frac{1}{2})(j_1+j_2+j_3)} |\mathbf{b}|^{1-3p}, \end{aligned}$$

using Proposition 2 with  $d = 3$ . Like in the previous cases, we can majorize the sum over domain D by an integral

$$\begin{aligned} \sum_{\mathbf{a} \in D_{\mathbf{j}}} |\mathbf{b}|^{-q(3p-1)} &= \sum_{\mathbf{a} \in D_{\mathbf{j}}} |2^{-2j_1}a_1^2 + 2^{-2j_2}a_2^2 + 2^{-2j_3}a_3^2|^{-\frac{1}{2}q(3p-1)} \\ &\lesssim 2^{j_1+j_2+j_3} \int_{|\tilde{\mathbf{a}}| \geq 2^{-j_3}\delta} |\tilde{\mathbf{a}}|^{-q(3p-1)} d\tilde{\mathbf{a}} \\ &\lesssim 2^{j_1+j_2} 2^{((3p-1)q-2)j_3}. \end{aligned}$$

In the second line, we have used  $2 - q(3p-1) < -1$ , which is a simple consequence of  $p > \frac{1}{q} + \frac{1}{2}$ . Putting things together for domain D, we obtain the bound

$$\begin{aligned} I_D &= \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{j_1 q} \left( \sum_{\mathbf{a} \in D_{\mathbf{j}}} |\langle \chi_{\mathbf{j}, \mathbf{a}} | \phi \rangle|^q \right) \\ &\lesssim \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{j_1 q} 2^{-q(p+\frac{1}{2})(j_1+j_2+j_3)} 2^{j_1+j_2} 2^{((3p-1)q-2)j_3} \\ &= \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{-qj_2} 2^{-\frac{q}{2}j_3} 2^{-(qp-\frac{q}{2}-1)(j_1-j_3)} 2^{-(qp-\frac{q}{2}-1)(j_2-j_3)} \\ &= \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{-\epsilon j_1} 2^{-(q+\epsilon)j_2} 2^{-(\frac{q}{2}-2\epsilon)j_3} 2^{-(qp-\frac{q}{2}-1-\epsilon)(j_1-j_3)} 2^{-(qp-\frac{q}{2}-1-\epsilon)(j_2-j_3)} \\ &\lesssim \sum_{j_1 \geq j_2 \geq j_3 \geq j_0} 2^{-\epsilon j_1} 2^{-(q+\epsilon)j_2} 2^{-(\frac{q}{2}-2\epsilon)j_3} \\ &< \infty, \end{aligned}$$

where an arbitrary parameter  $0 < \epsilon < \min\{pq-1-\frac{q}{2}, \frac{q}{4}\}$  has been introduced in order to demonstrate convergence in the norm (1.9). The definition of  $\epsilon$  is justified by  $pq-1-\frac{q}{2} > 0$ .  $\square$

**Remark 1.** In our proof we have only assumed the minimal number of vanishing moments  $p > \alpha + 1$  for the univariate wavelet basis that are required by the norm equivalence (1.4). This is however not necessary, instead it is possible to assume a larger number of vanishing moments, which would slightly simplify parts of the proof.

The method used for the proof of Lemma 1 can be applied in a straightforward manner to best  $N$ -term approximation spaces for isotropic wavelets (1.5).

**Corollary 1.** *Suppose  $\phi$  satisfies the assumptions of Lemma 1. Then  $\phi \in B_q^\alpha(L^q(\Omega))$  for all  $q > 0$  and  $\alpha = \frac{3}{q} - \frac{1}{2}$ .*

*Proof.* Due to the norm equivalence (1.7) for isotropic wavelets, the corollary is equivalent to

$$\sum_{j \geq j_0} \sum_{\mathbf{s}} 2^{jq} \left( \sum_{\mathbf{a}} \left| \langle \gamma_{j,\mathbf{a}}^{(\mathbf{s})} | \phi \rangle \right|^q \right) < \infty.$$

The proof of this statement resembles to our discussion of domain A in Lemma 1. A parameter  $L$  is chosen in such a way that  $\text{supp } \psi^{(s)} \subset [-L/2, L/2]$ . We decompose the sum into two parts and estimate them separately. In the first case with  $\text{dist}(\text{supp } \gamma_{j,\mathbf{a}}^{(s)}, \mathbf{0}) > 2^{-j}L$ , we get

$$\begin{aligned} \sum_{j \geq j_0} \sum_{\mathbf{s}} 2^{jq} \sum_{|\mathbf{a}| > 2L} \left| \langle \gamma_{j,\mathbf{a}}^{(\mathbf{s})} | \phi \rangle \right|^q &\lesssim \sum_{j \geq j_0} 2^{-\frac{3}{2}jq} \sum_{|\mathbf{a}| > 2L} |\mathbf{a}|^{q(1-p)} \\ &\lesssim \sum_{j \geq j_0} 2^{-\frac{3}{2}jq} \int_{2L}^{2^j \text{diam}(\Omega)} |\mathbf{a}|^{q(1-p)} d\mathbf{a} \\ &\lesssim \sum_{j \geq j_0} \begin{cases} 2^{-(qp + \frac{1}{2}q - 3)j}, & \text{if } -1 < q(1-p) + 2 < -1 + \frac{3}{2}q \\ (j+1)2^{-\frac{3}{2}jq}, & \text{if } q(1-p) + 2 = -1 \\ 2^{-\frac{3}{2}jq}, & \text{if } q(1-p) + 2 < -1 \end{cases} \\ &< \infty, \end{aligned}$$

where we have used  $p > \alpha = \frac{3}{q} - \frac{1}{2}$  vanishing moments of the univariate wavelet in one direction only. The remainder can be directly estimated

$$\sum_{j \geq j_0} \sum_{\mathbf{s}} 2^{jq} \sum_{|\mathbf{a}| \leq 2L} \left| \langle \gamma_{j,\mathbf{a}}^{(\mathbf{s})} | \phi \rangle \right|^q \lesssim \sum_{j \geq j_0} 2^{-\frac{3}{2}jq} < \infty. \quad \square$$

## 2.2. Optimal convergence rates for wavelet approximations

Under the rather mild assumption of asymptotic smoothness near electron-nuclear cusps, Lemma 1 shows that for a univariate wavelet basis  $\psi_{j,a}^{(s)}$  with  $p > \alpha + 1$  vanishing moments and  $\psi_{j,a} \in B_q^\beta(L^q(I))$  for some  $\beta > \alpha + 1$ , the orbitals  $\phi_i$  belong to the approximation spaces  $A_q^\alpha(H^1(\Omega))$  for the corresponding anisotropic wavelet basis  $\chi_{j,\mathbf{a}}$ . This means that it is possible to achieve convergence rates  $\sigma_N(\phi) \sim N^{-\alpha}$  for the approximation error in  $H^1(\Omega)$ . For comparison, the corresponding isotropic wavelet basis  $\gamma_{j,\mathbf{a}}^{(s)}$  can only achieve convergence rates  $\sigma_N(\phi) \sim N^{-\alpha/3}$  because of the dimensional dependence in equation (1.6).

In order to extend our discussion to one-electron reduced density matrices  $\rho(\mathbf{x}, \mathbf{y})$ , we state the following corollary which is an immediate consequence of the proof of Lemma 1.

**Corollary 2.** *Suppose that  $\rho \in C^\infty((\mathbb{R}^3 \setminus \mathbf{R}_a) \times (\mathbb{R}^3 \setminus \mathbf{R}_b))$  is asymptotically smooth*

$$|\partial_{\mathbf{x}}^\beta \partial_{\mathbf{y}}^\gamma \rho(\mathbf{x}, \mathbf{y})| \leq c_{\beta,\gamma} |\mathbf{x} - \mathbf{R}_a|^{1-|\beta|} |\mathbf{y} - \mathbf{R}_b|^{1-|\gamma|}, \quad \mathbf{x} \neq \mathbf{R}_a, \mathbf{y} \neq \mathbf{R}_b \text{ and } |\beta|, |\gamma| \geq 1$$

*in a bounded neighbourhood  $\Omega_{\mathbf{R}_a} \times \Omega_{\mathbf{R}_b} \subset \mathbb{R}^3 \times \mathbb{R}^3$  of two, possibly identical, points  $\mathbf{R}_a, \mathbf{R}_b$ . Then  $\rho \in \mathcal{B}_q^\alpha(\Omega_{\mathbf{R}_a} \times \Omega_{\mathbf{R}_b})$  for all  $q > 0$  and  $\alpha = \frac{1}{q} - \frac{1}{2}$ .*

It is obvious that this assumption is fulfilled for HF and DFT density matrices if the orbitals itself are asymptotically smooth. Therefore, we can achieve the same convergence rates  $\sigma_N(\rho) \sim N^{-\alpha}$  as for orbitals irrespective of the increase of dimension.

We want to conclude this section with a brief discussion of related results for GTO basis sets

$$\phi_{lm}^{\text{GTO}}(r, \varphi, \theta) = C_{lm} Y_{lm}(\varphi, \theta) r^l e^{-\zeta r^2},$$

which are by far the most popular basis sets in quantum chemistry [15]. Despite their tremendous success in applications, not much rigorous results concerning their approximation properties have been reported in the literature [1, 21]. This work focuses on the approximation of single STO basis functions (2.1) by linear combinations of GTOs, where almost exponential convergence rates  $\sigma_N(\phi_{nlm}^{\text{STO}}) \sim e^{-\mu\sqrt{N}}$  can be achieved. Whether this result remains valid for molecules seems to be an open issue, although there exists some numerical evidence in favor of it [13]. We want to mention, that our assumption of asymptotic smoothness was motivated by STO basis sets but is more general and does not rely on spherical symmetry. Our results, therefore, apply to systems with several nuclei as well. Furthermore GTO basis sets are not stable in a sense that there is no norm equivalence of an appropriate Hilbert space to a weighted  $\ell_2$  space of coefficients. The question remains open, whether wavelets provide a serious alternative to GTOs in realistic HF and DFT calculations. Despite some remarkable successes concerning efficient implementations for these methods [14, 27, 28], it is hard to compete for wavelets with GTO based programs which have been developed and optimized over the last three decades [15]. Finally we want mention that there exists a variety of applications for HF and DFT methods where GTO basis sets are not really appropriate due to geometrical constraints, like for quasi two-dimensional many-particle systems which appear *e.g.* in semiconductor heterostructures. For such kind of systems wavelets seem to be an interesting alternative [9].

*Acknowledgements.* This work was supported by the Deutsche Forschungsgemeinschaft (SPP 1145).

## REFERENCES

- [1] D. Braess, Asymptotics for the approximation of wave functions by exponential sums. *J. Approx. Theory* **83** (1995) 93–103.
- [2] H.-J. Bungartz and M. Griebel, Sparse grids. *Acta Numerica* **13** (2004) 147–269.
- [3] A. Cohen, R.A. DeVore and R. Hochmuth, Restricted nonlinear approximation. *Constr. Approx.* **16** (2000) 85–113.
- [4] R.A. DeVore, Nonlinear approximation. *Acta Numerica* **7** (1998) 51–150.
- [5] R.A. DeVore, B. Jawerth and V. Popov, Compression of wavelet decompositions. *Amer. J. Math.* **114** (1992) 737–785.
- [6] R.A. DeVore, S.V. Konyagin and V.N. Temlyakov, Hyperbolic wavelet approximation. *Constr. Approx.* **14** (1998) 1–26.
- [7] H.-J. Flad, W. Hackbusch, D. Kolb and R. Schneider, Wavelet approximation of correlated wavefunctions. I. Basics. *J. Chem. Phys.* **116** (2002) 9641–9657.
- [8] H.-J. Flad, W. Hackbusch, H. Luo and D. Kolb, Diagrammatic multiresolution analysis for electron correlations. *Phys. Rev. B* **71** (2005) 125115.
- [9] H.-J. Flad, W. Hackbusch, H. Luo and D. Kolb, Wavelet approach to quasi two-dimensional extended many-particle systems. I. supercell Hartree-Fock method. *J. Comp. Phys.* **205** (2005) 540–566.
- [10] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Ostergaard Sorensen, On the regularity of the density of electronic wavefunctions. *Contemp. Math.* **307** (2002) 143–148.
- [11] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Ostergaard Sorensen, The electron density is smooth away from the nuclei. *Commun. Math. Phys.* **228** (2002) 401–415.
- [12] J. Garcke and M. Griebel, On the computation of the eigenproblems of hydrogen and helium in strong magnetic and electric fields with the sparse grid combination technique. *J. Comp. Phys.* **165** (2000) 694–716.
- [13] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper and J. Olsen, Basis-set convergence of the energy in molecular Hartree-Fock calculations. *Chem. Phys. Lett.* **302** (1999) 437–446.
- [14] R.J. Harrison, G.I. Fann, T. Yanai, Z. Gan and G. Beylkin, Multiresolution quantum chemistry: Basic theory and initial applications. *J. Chem. Phys.* **121** (2004) 11587–11598.
- [15] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory*, Wiley, New York (1999).
- [16] R.N. Hill, Rates of convergence and error estimation formulas for the Rayleigh-Ritz variational method. *J. Chem. Phys.* **83** (1985) 1173–1196.
- [17] M. Hoffmann-Ostenhof and R. Seiler, Cusp conditions for eigenfunctions of n-electron systems, *Phys. Rev. A* **23** (1981) 21–23.
- [18] M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and H. Stremnitzer, Local properties of Coulombic wave functions. *Commun. Math. Phys.* **163** (1994) 185–215.
- [19] M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Ostergaard Sorensen, Electron wavefunctions and densities for atoms. *Ann. Henri Poincaré* **2** (2001) 77–100.

- [20] T. Kato, On the eigenfunctions of many-particle systems in quantum mechanics. *Commun. Pure Appl. Math.* **10** (1957) 151–177.
- [21] W. Kutzelnigg, Theory of the expansion of wave functions in a Gaussian basis. *Int. J. Quantum Chem.* **51** (1994) 447–463.
- [22] W. Kutzelnigg and J.D. Morgan III, Rates of convergence of the partial-wave expansions of atomic correlation energies. *J. Chem. Phys.* **96** (1992) 4484–4508.
- [23] E.H. Lieb and B. Simon, The Hartree-Fock theory for Coulomb systems. *Commun. Math. Phys.* **53** (1977) 185–194.
- [24] H. Luo, D. Kolb, H.-J. Flad, W. Hackbusch and T. Koprucki, Wavelet approximation of correlated wavefunctions. II. Hyperbolic wavelets and adaptive approximation schemes. *J. Chem. Phys.* **117** (2002) 3625–3638.
- [25] P.-A. Nitsche, *Best  $N$ -term approximation spaces for sparse grids*, Research Report No. 2003-11, Seminar für Angewandte Mathematik, ETH Zürich.
- [26] R. Schneider, *Multiskalen- und Wavelet-Matrixkompression*, Teubner, Stuttgart (1998).
- [27] T. Yanai, G.I. Fann, Z. Gan, R.J. Harrison and G. Beylkin, Multiresolution quantum chemistry in multiwavelet basis: Hartree-Fock exchange. *J. Chem. Phys.* **121** (2004) 6680–6688.
- [28] T. Yanai, G.I. Fann, Z. Gan, R.J. Harrison and G. Beylkin, Multiresolution quantum chemistry in multiwavelet basis: Analytic derivatives for Hartree-Fock and density functional theory. *J. Chem. Phys.* **121** (2004) 2866–2876.
- [29] H. Yserentant, On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives. *Numer. Math.* **98** (2004) 731–759.
- [30] H. Yserentant, Sparse grid spaces for the numerical solution of the electronic Schrödinger equation. *Numer. Math.* **101** (2005) 381–389.