s^*-compressibility of discrete Hartree-Fock equations

Heinz-Jürgen Flad and Reinhold Schneider

Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, D-10623 Berlin

October 17, 2008

Abstract

The Hartree-Fock equations are widely accepted as the basic model of electronic structure calculation which serves as a canonical starting point for more sophisticated many-particle methods. We have studied the s^*-*compressibility* for Galerkin discretizations of the Hartree-Fock equations in wavelet bases. Our focus is on the compression of Galerkin matrices from nuclear Coulomb potentials and nonlinear terms in the Fock operator which hitherto has not been discussed in the literature. It can be shown that the s^*-compressibility is in accordance with convergence rates obtained from best N-term approximation for solutions of the Hartree-Fock equations. This is a necessary requirement in order to achieve numerical solutions for these equations with optimal complexity using the recently developed adaptive wavelet algorithms of Cohen, Dahmen and DeVore.

Keywords: Hartree-Fock equations, matrix compression, best N-term approximation

1 Introduction

1.1 The Hartree-Fock model in a nutshell

Many important applications of quantum theory in chemistry can be traced back to solutions of the stationary electronic Schrödinger equation within the Born-Oppenheimer approximation, neglecting relativistic effects. The Schrödinger equation for a system of N electrons exposed to the stationary electrostatic field of K nuclei with charges $\{Z_i\}_{i=1,K}$ corresponds to a symmetry restricted eigenvalue problem

$$H\Psi_i = E_i\Psi_i \text{ with } \Psi_i \in H^1_e(\mathbb{R}^{3N}) \times \{\pm 1/2\}^N \tag{1.1}$$

with respect to the nonrelativistic electronic Hamiltonian

$$H = -\frac{1}{2}\sum_{i=1}^{N}\Delta_i \; - \; \sum_{k=1}^{K}\sum_{i=1}^{N}\frac{Z_k}{|\mathbf{x}_i - \mathbf{A}_k|} \; + \; \sum_{i<j}^{N}\frac{1}{|\mathbf{x}_i - \mathbf{x}_j|}. \tag{1.2}$$

The symmetry constraint that enters via Pauli's principle can be most easily expressed by introducing spin degrees of freedom into the electron coordinates, i.e. $\underline{\mathbf{x}}_i := (\mathbf{x}_i, \xi_i)$ with $\mathbf{x}_i \in \mathbb{R}^3$ and $\xi_i \in \{\pm 1/2\}$. In this case Pauli's principle requires the wavefunction to be antisymmetric with respect to the permutation of two electron coordinates, i.e.,

$$\Psi_i(\ldots, \underline{\mathbf{x}}_i, \ldots, \underline{\mathbf{x}}_j, \ldots) = -\Psi_i(\ldots, \underline{\mathbf{x}}_j, \ldots, \underline{\mathbf{x}}_i, \ldots).$$

Approximate many-electron wavefunctions can be obtained by taking the exterior tensor product

$$\det[\phi_i(\underline{\mathbf{x}}_j)]^{j=1,\ldots,N}_{i\in(i_1,\ldots i_N)} := \frac{1}{\sqrt{N!}}\,\phi_{i_1} \wedge \phi_{i_2} \wedge \ldots \wedge \phi_{i_N},$$

of N orthonormal single-electron wavefunctions $\{\phi_i\}_{i\in(i_1,\ldots i_N)}$ with $\phi_i \in H^1(\mathbb{R}^3) \times \{\pm 1/2\}$. In quantum chemistry, these exterior products are denoted as Slater determinants. The nonlinear subspace of all N-electron Slater determinants

$$\Gamma_N := \left\{ \Phi \mid \Phi := \det[\phi_i(\underline{\mathbf{x}}_j)]_{i\in(i_1,\ldots i_N)}^{j=1,\ldots,N} \text{ with } \phi_i \in H^1(\mathbb{R}^3) \times \{\pm 1/2\} \text{ and } \langle \phi_i | \phi_j \rangle = \delta_{i,j} \right\}$$

provides the basis for the Hartree-Fock model which is defined as a rank-1 approximation for the many-electron wavefunction which minimizes the variational energy

$$E_0 = \inf_{\Phi\in\Gamma_N} \frac{\langle \Phi | H\Phi \rangle}{\langle \Phi | \Phi \rangle}. \tag{1.3}$$

Existence of a minimizer has been demonstrated in [24, 25], however no uniqueness results are presently known. For notational simplicity we restrict ourselves to the closed shell Hartree-Fock model, where the eigenfunctions are assumed to be real functions of the form $\phi_i(\mathbf{x})\chi(\xi)$ with the same spatial part for both spins $\{\pm 1/2\}$. The variational problem (1.3) leads to the nonlinear Hartree-Fock equations

$$\mathfrak{h}\phi_i = \varepsilon_i \, \phi_i, \tag{1.4}$$

with single-particle Hamiltonian

$$\mathfrak{h} = -\frac{1}{2}\Delta - \sum_{k=1}^{K} \frac{Z_k}{|\mathbf{x} - \mathbf{A}_k|} + V_H(\mathbf{x}) - U, \tag{1.5}$$

the so-called Fock operator, which contains the Hartree potential

$$V_H(\mathbf{x}) = 2\sum_{i=1}^{N/2} \int_{\mathbb{R}^3} \frac{|\phi_i(\mathbf{y})|^2}{|\mathbf{x} - \mathbf{y}|} d^3y, \tag{1.6}$$

and the nonlocal exchange operator

$$Uv(\mathbf{x}) = \sum_{i=1}^{N/2} \int_{\mathbb{R}^3} \frac{\phi_i(\mathbf{x})\phi_i(\mathbf{y})v(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d^3y, \tag{1.7}$$

where only those eigenfunctions enter into the nonlinear terms (1.6) and (1.7) of the Fock operator which belong to the lowest $N/2$ eigenvalues.

1.2 Galerkin discretization of the Hartree-Fock equations

Gaussian-type atomic centered basis functions [21] are by far the most popular basis sets in electronic structure calculations. Amongst others, Gaussian-type basis functions have the significant advantage that all integrals required for a Galerkin discretization of the Fock operator can be evaluated analytically. Despite their tremendous success in applications, only little is known about their approximation properties [4, 23]. This is mainly due to the fact that these bases represent essentially overcomplete frames which are best described in terms of a nonlinear approximation theory. Another related obstacle for a rigorous analysis is that they are not stable in the sense of a Riesz basis in our envisaged Hilbert spaces L^2 and H^1. Alternative discretization schemes for the Hartree-Fock equations based on finite differences [22, 1] or finite elements [19, 3] have been rarely studied in the literature. Nevertheless these methods achieved unsurpassed accuracies for linear molecules and served as benchmarks for the development of Gaussian basis sets. Within the last few years wavelets attracted considerable interest in electronic structure calculations. Most of this work has been done in the context of density functional theory [2, 12, 26, 16, 20]. However Yanai et al. [33] report very accurate Hartree-Fock calculations for some small molecules. A significant adavantage of wavelets over Gaussian-type basis functions is that they provide stable basis sets which enable the adaptive approximation of singularities and compression schemes for Galerkin discretizations of a large class of Calderón-Zygmund and pseudo-differential operators [32]. Within the present work we want to study compression schemes for wavelet Galerkin discretizations of the Fock operator. This is a necessary prerequisite for the adaptive wavelet algorithms of Cohen, Dahmen and DeVore [8]. The Fock operator resembles to a pseudo-differential operator, however, due to the presence of additional point like Coulomb singularities at the nuclei, the symbols do not belong to the standard classes which requires an extended calculus for pseudo-differential operators on manifolds with conical singularities, cf. [11]. As a consequence the compression schemes developed e.g. in Ref. [32] are not immediately applicable and require further adaptive refinement schemes which have been studied within the present work.

1.3 Tensor products of the Fock operator in post Hartree-Fock models

On a first glance the Hartree-Fock model seems to be a rather crude approximation for many-electron systems. However, in many applications it turned out to be amazingly accurate. This is far from obvious from a mathematical point of view and it was only recently that the work of Yserentant shed some light on the remarkable success of the Hartree-Fock model. It was proven by Yserentant [34] that solutions of the many-particle Schrödinger equation (1.1) belong to Sobolev spaces of mixed partial derivatives. This is the basic regularity condition for sparse grids or hyperbolic cross approximations [5] which enable low-rank tensor product approximations for generically high-dimensional many-particle problems. In this sense Hartree-Fock can be considered as the best rank-one approximation with respect to the total energy. Despite its amazing accuracy, the limitations of the Hartree-Fock model become clear if one considers the small energy differences which are of relevance for chemical phenomena. Let us just mention the binding energy of a certain chemical bond which typically represents only a small fraction of the total energy of a molecule. In order to achieve chemical accuracy it is therefore necessary to go beyond Hartree-Fock and to consider more sophisticated many-particle methods. Nevertheless, the Hartree-Fock model remains an essential ingredient of these methods. In a more formal sense this can be seen from the appearance of tensor product Fock operators which enter into the equations. We just want to mention Møller-Plesset perturbation theory [21], where the Hamiltonian (1.2) is split up into

$$H = H_0 + H_1 \quad \text{with } H_0 = \sum_{i=1}^{N} \mathfrak{h}_i,$$

where $\mathfrak{h}_i$, with $i = 1, \ldots, N$, are Fock operators acting on the variables $\mathbf{x}_i$. Actually, the Hamiltonian H_0 has to be considered as a sum of tensor products $I_1 \otimes \cdots \otimes \mathfrak{h}_i \otimes \cdots \otimes I_N$ of a Fock operator and identities. Second-order Møller-Plesset perturbation theory can be written as a system of integro-differential equations [28, 29] of the form

$$(\mathfrak{h}_1 + \mathfrak{h}_2 - \epsilon_a - \epsilon_b)\, \tau^{\pm}_{a,b}(\mathbf{x}_1, \mathbf{x}_2) = -\mathfrak{Q} \frac{1}{|\mathbf{x}_1 - \mathbf{x}_2|} \left(\phi_a(\mathbf{x}_1)\phi_b(\mathbf{x}_2) \pm \phi_b(\mathbf{x}_1)\phi_a(\mathbf{x}_2) \right),$$
$$\mathfrak{Q}\tau^{\pm}_{a,b}(\mathbf{x}_1, \mathbf{x}_2) = \tau^{\pm}_{a,b}(\mathbf{x}_1, \mathbf{x}_2),$$

where for each pair of occupied Hartree-Fock orbitals ϕ_a, ϕ_b different pair-correlation functions $\tau^{\pm}_{a,b}$ with respect to singlet (+) and triplet (−) states exist. The operator $\mathfrak{Q}$ is a projection operator on the virtual two-particle space i.e. $\mathfrak{Q} := (1 - \mathfrak{q}_1)(1 - \mathfrak{q}_2)$ with $\mathfrak{q} := \sum_{i=1}^{N/2} |\phi_i\rangle\langle\phi_i|$.

Our brief discussion clearly demonstrates that the Galerkin discretization of the Fock operator and its compressibility is also a central issue for post Hartree-Fock methods. Results on the s^*-compressibility are discussed in a subsequent publication.

2 s^*-compressibility of the Fock matrix

2.1 Some technical prerequisites

Within the present work we focus on compression schemes for the Galerkin discretization of the Fock operator which enable the application of adaptive wavelet algorithms for the solution of the Hartree-Fock equations with optimal computational complexity. The adaptive wavelet algorithm for linear equations from Ref. [8] has been recently generalized to eigenvalue problems [9, 27] which can be applied within the self-consistent iteration schemes usually employed for the numerical solution of the Hartree-Fock equations. These algorithms rest on two pillars, one is the compression of Galerkin matrices and the other is the best N-term approximation of solutions. In our previous work we have studied best N-term approximation spaces for solutions of the Hartree-Fock equations [14]. The purpose of the present work is to demonstrate the existence of compression schemes with error estimates compatible with the convergence rates obtained from best N-term approximation theory. It is obvious that different parts of the Fock operator (1.5) require different compression schemes, i.e. the Laplacian, the various local singular potentials and the nonlocal exchange operator have to be considered separately. Although a lot of work has been already devoted to study the compressibility of Galerkin matrices for a large class of operators, electronic structure calculations pose an additional difficulty due to the presence of various kinds of singularities in the Fock operator (1.5) at the locations of the nuclei. Compression schemes for typical integral operators so far considered in the mathematical literature [32] take into account only diagonal singularities.

In the following we consider the Galerkin discretization of the Fock operator in a isotropic 3d-wavelet basis

$$\begin{aligned} \gamma^{(1)}_{j,\mathbf{a}}(\mathbf{x}) &= \psi_{j,a_1}(x_1)\,\varphi_{j,a_2}(x_2)\,\varphi_{j,a_3}(x_3), \\ &\vdots \\ \gamma^{(4)}_{j,\mathbf{a}}(\mathbf{x}) &= \psi_{j,a_1}(x_1)\,\psi_{j,a_2}(x_2)\,\varphi_{j,a_3}(x_3), \\ &\vdots \\ \gamma^{(7)}_{j,\mathbf{a}}(\mathbf{x}) &= \psi_{j,a_1}(x_1)\,\psi_{j,a_2}(x_2)\,\psi_{j,a_3}(x_3), \end{aligned} \tag{2.1}$$

which correspond to tensor products of univariate wavelets $\psi_{j,a}(x) = 2^{\frac{1}{2}j}\psi(2^{-j}x - a)$ and scaling functions $\varphi_{j,a}(x) = 2^{\frac{1}{2}j}\varphi(2^{-j}x - a)$ on the same level j. We assume that the univariate wavelets represent Riesz bases in $L_2(\mathbb{R})$ and, after rescaling, in $H^1(\mathbb{R})$. According to our definition, the 3d-wavelets have a different number of vanishing moments, r, $2r$ or $3r$, depending on the number of univariate wavelets, with r vanishing moments, in the tensor product. In the remaining part of the paper we skip the superscripts $\gamma^{(1)\cdots(7)}$, which specify the type of 3d-wavelet, in our notation. Instead, w.l.o.g. the following estimates assume only the minimal number r of vanishing moments. Within the present work, we require the univariate wavelets to have compact support and satisfy the common regularity assumption [31], i.e. a Bernstein inequality

$$\|\gamma_{j,\mathbf{a}}\|_{W^t_\infty} \lesssim 2^{(\frac{3}{2}+t)j} \quad \text{for } t \in [0, s], \tag{2.2}$$

where $r \geq s+1$ has been assumed. This assumption is typically satisfied e.g. for univariate spline wavelets. Here and in the following $a \lesssim b$ means that a is uniformly bounded by some constant multiple of b. Similarly $a \simeq b$ means that the quantities can be uniformly bounded by some constant multiple of each other.

In order to obtain our basic estimates we will frequently use the following propositions.

Proposition 1. *Given the function $f(\mathbf{x})$ with $\mathbf{x} \in \mathbb{R}^3$ and an isotropic 3d-wavelet $\gamma_{j,\mathbf{a}}$ with $r \geq s+1$ vanishing moments in the i'th direction and which satisfies the Bernstein inequality (2.2).*

(i) Suppose f is Lipschitz continuous, i.e. $f \in C^{0,1}(\mathbb{R}^3)$, then

$$\left| \int\limits_{\mathbb{R}^3} f(\mathbf{x})\, \gamma_{j,\mathbf{a}}(\mathbf{x}) d^3x \right| \lesssim 2^{-\frac{5}{2}j} \|\partial_i f\|_{L_\infty(\operatorname{supp}\gamma_{j,\mathbf{a}})}.$$

(ii) Suppose f belongs to $W^t_\infty(\operatorname{supp}\gamma_{j,\mathbf{a}})$ for $0 \le t \le r$, then

$$\begin{aligned} \left| \int\limits_{\mathbb{R}^3} f(\mathbf{x})\, \gamma_{j,\mathbf{a}}(\mathbf{x}) d^3x \right| &\lesssim 2^{-(t+\frac{3}{2})j} \|\partial_i^t f\|_{L_\infty(\operatorname{supp}\gamma_{j,\mathbf{a}})} && (2.3)\\ &\le 2^{-(t+\frac{3}{2})j} \|f\|_{W^t_\infty(\operatorname{supp}\gamma_{j,\mathbf{a}})}. && (2.4) \end{aligned}$$

Proof. The proof is a direct consequence of the vanishing moment property and Taylor's theorem using the integral representation for the remainder of the Taylor series, see e.g. [14] for details. □

Proposition 2. *Suppose f has compact support Ω and belongs to $W^t_\infty(\mathbb{R}^3)$, with $t \ge 1$, then*

$$\left| \partial_i^{t+1} \int\limits_\Omega \frac{f(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} d^3y \right| \lesssim \|\partial_i^t f\|_{L_\infty(\Omega)} \int\limits_\Omega \frac{1}{|\mathbf{x}-\mathbf{y}|^2} d^3y. \qquad (2.5)$$

Proof. From Poisson's equation and $f \in W^t_\infty(\mathbb{R}^3)$ it follows that in a distributional sense

$$\partial_i^t f(\mathbf{x}) = \partial_i^t \Delta \int\limits_\Omega \frac{f(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} d^3y = \Delta \partial_i^t \int\limits_\Omega \frac{f(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} d^3y,$$

and therefore

$$\partial_i^t \int\limits_\Omega \frac{f(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} d^3y = \int\limits_\Omega \frac{\partial_i^t f(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} d^3y.$$

Since $\partial_i^t f \in L_\infty(\Omega)$ the left side actually belongs to $C^1(\mathbb{R}^3)$ and we can differentiate once more according to Lemma 4.1 in Ref. [18] to get the estimate (2.5). □

Proposition 3. *Suppose f belongs to $W^t_\infty(\Omega)$, with $t \ge 1$, for a Lipschitz domain Ω. Then the t'th order derivatives of the Coulomb potential satisfy the estimates*

$$\left| \partial_i^t \int\limits_\Omega \frac{f(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} d^3y \right| \lesssim \|\partial_i^{t-1} f\|_{L_\infty(\Omega)} \int\limits_\Omega \frac{1}{|\mathbf{x}-\mathbf{y}|^2} d^3y. \qquad (2.6)$$

Furthermore for $\mathbf{x} \in \Omega$, the $t+1$'th order derivatives satisfy the following estimates

$$\left| \partial_i^{t+1} \int\limits_\Omega \frac{f(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} d^3y \right| \lesssim \|\partial_i^t f\|_{L_\infty(\Omega)} \int\limits_\Omega \frac{1}{|\mathbf{x}-\mathbf{y}|^2} d^3y + \|\partial_i^{t-1} f\|_{L_\infty(\Omega)} \int\limits_{\partial\Omega} \frac{1}{|\mathbf{x}-\mathbf{y}|^2} ds \qquad (2.7)$$

Proof. We have $f \in C^{t-1,1}(\overline{\Omega})$ because of $f \in W^t_\infty(\Omega)$, and therefore

$$\partial_i^{t-1} \int\limits_\Omega \frac{f(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} d^3y = \int\limits_\Omega \frac{\partial_i^{t-1} f(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} d^3y.$$

The estimate (2.6) follows from Lemma 4.1 in Ref. [18]. According to Lemma 4.2 in Ref. [18], we can actually differentiate twice and obtain for $\mathbf{x} \in \Omega$ the estimate

$$\left| \partial_i^2 \int\limits_\Omega \frac{\partial_i^{t-1} f(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} d^3y \right| \lesssim \int\limits_\Omega \frac{1}{|\mathbf{x}-\mathbf{y}|^3} |\partial_i^{t-1} f(\mathbf{x}) - \partial_i^{t-1} f(\mathbf{y})| d^3y + \|\partial_i^{t-1} f\|_{L_\infty(\Omega)} \int\limits_{\partial\Omega} \frac{1}{|\mathbf{x}-\mathbf{y}|^2} ds.$$

Since $\partial_i^{t-1} f \in C^{0,1}(\overline{\Omega})$, we get for $\mathbf{x}, \mathbf{y} \in \Omega$

$$\frac{\left|\partial_i^{t-1} f(\mathbf{x}) - \partial_i^{t-1} f(\mathbf{y})\right|}{|\mathbf{x}-\mathbf{y}|} \lesssim \left\|\partial_i^t f\right\|_{L_\infty(\Omega)},$$

from which (2.7) follows. □

A basic prerequisite of adaptive wavelet algorithms is the s^*-*compressibility* of the Galerkin matrix obtained from the discretization of the Fock operator (1.5) in the isotropic 3d-wavelet basis (2.1). Various equivalent definitions of s^*-compressibility have been given in the literature, cf. [8, 30, 31], we adopted the definition from Ref. [31]:

Definition 1. *An infinite matrix* $\mathbf{A}$ *is called* s^**-compressible if for each* $n \in \mathbb{N}$ *an approximate matrix* $\mathbf{A}^{(n)}$ *can be constructed, that has in each row and column at most* $O(2^n)$ *nonzero entries, which satisfies the spectral norm estimate* $\|\mathbf{A} - \mathbf{A}^{(n)}\| \lesssim 2^{-\alpha n}$ *for any* $0 < \alpha < s^*$.

In order to derive spectral norm estimates we use the following version of Schur's Lemma, cf. Ref. [32].

Lemma 1. *Given a symmetric infinite Galerkin matrix* $\mathbf{A} := \{A_{(j,\mathbf{a}),(l,\mathbf{b})}\}_{(j,\mathbf{a}),(l,\mathbf{b})\in\Lambda}$ *with respect to a wavelet basis with index set* Λ. *The spectral norm of the Galerkin matrix can be estimated according to*

$$\|\mathbf{A}\| \lesssim \sup_{(j,\mathbf{a})\in\Lambda} \sum_{(l,\mathbf{b})\in\Lambda} 2^{p(j-l)} \left|A_{(j,\mathbf{a}),(l,\mathbf{b})}\right|, \tag{2.8}$$

for any $p \in \mathbb{R}$.

The presence of nonlinear Hartree and exchange terms in the Fock operator (1.5) requires special considerations concerning the regularity of the solutions. As already mentioned before, we consider an iterative solution scheme for the Hartree-Fock equations. This means that in each iteration step a nonlinear eigenvalue problem has to be solved where the operator depends, however, on the solutions of the previous iteration. Various iteration schemes have been proposed in the quantum chemistry literature [6] but only for the level-shifting algorithm convergence has been actually proven. The level-shifting algorithm follows the iterative scheme

$$\cdots \longrightarrow \{\phi_n^{(i)}\}_{n=1,\ldots,N/2} \longrightarrow \mathfrak{h}_{\text{shift}}^{(i)} := \mathfrak{h}^{(i)} - b\mathcal{P}^{(i)} \longrightarrow \{\phi_n^{(i+1)}\}_{n=1,\ldots,N/2} \longrightarrow \cdots, \tag{2.9}$$

where in step i, the eigenfunctions $\{\phi_n^{(i)}\}_{n=1,\ldots,N/2}$ which belong to the lowest $N/2$ eigenvalues of a shifted Fock operator $\mathfrak{h}_{\text{shift}}^{(i-1)}$ are selected accordine to the aufbau principle in order to construct a new shifted Fock operator $\mathfrak{h}_{\text{shift}}^{(i)}$. Here the operator $\mathcal{P}^{(i)}$ denotes an orthogonal projection on the space spanned by the eigenfunctions $\{\phi_n^{(i)}\}_{n=1,\ldots,N/2}$. The convergence of this algorithm for a sufficiently large parameter b, to "self-consistent-field" solutions of the Hartree-Fock equations (1.4) has been proven by Cancès and Le Bris [7]. In a previous paper [15] we have studied the asymptotic behaviour of solutions within such an iteration scheme and of the final "self-consistent-field" solutions of the Hartree-Fock equations. For the present work we require all intermediate and final solutions to be *asymptotically smooth* in a neighbourhood of a nucleus, i.e.

$$|\partial_{\mathbf{x}}^{\beta} \phi_n(\mathbf{x})| \leq C_\beta |\mathbf{x} - \mathbf{A}_k|^{1-|\beta|}, \text{ for } \mathbf{x} \neq \mathbf{A}_k, \quad \text{and } |\beta| \geq 1. \tag{2.10}$$

This property is an immediate consequence of the following theorem [15]:

Theorem 1. *All intermediate iterative solutions* $\{\phi_n^{(i)}\}_{n=1,\dots,N/2}$ *as well as the final self-consistent-field solutions* $\{\phi_n\}_{n=1,\dots,N/2}$ *of the Hartree-Fock equations, obtained via the level-shifting algorithm, exhibit Taylor asymptotics (polar coordinates) in a neighbourhood* Ω_k *of any of the* $k = 1,\dots K$ *nuclei, i.e.*

$$\phi = \omega(r)\sum_{j=0}^{l-1} c_j(\varphi,\theta)r^j \; + \; \Phi_l,$$

with $\Phi_l \in C_B^m(\Omega_k)$ *for* $l > m$*, provided that the initial guess* $\{\phi_n^{(0)}\}_{n=1,\dots,N/2}$ *possesses this property. Here* $\omega \in C_0^\infty(\Omega_k)$ *denotes a cut-off function which is equal to 1 in a neighbourhood of the nucleus. Furthermore we have* $\phi \in \mathcal{S}(\mathbb{R}^3 \setminus \cup_{k=1}^K \Omega_k)$.

A further consequence of this theorem is the asymptotic smoothness of the Hartree-potential in a neighbourhood of a nucleus, i.e.

$$\left|\partial_{\mathbf{x}}^\beta V_H(\mathbf{x})\right| \le C_\beta |\mathbf{x} - \mathbf{A}|^{3-|\beta|} \text{ for } \; \mathbf{x} \neq \mathbf{A}_k \; \text{ and } \; |\beta| \geq 3,$$

which is an immediate consequence of the asymptotic expansion of the theorem, cf. [15] for further details. The Hartree and Coulomb potentials of the nuclei can therefore be combined into the Coulomb potential

$$V(\mathbf{x}) = -\sum_{k=1}^K \frac{Z_k}{|\mathbf{x} - \mathbf{A}_k|} + V_H(\mathbf{x}), \tag{2.11}$$

which represents the electrostatic potential created by the averaged electron density and the nuclei. This combined Coulomb potential satisfies the asymptotic smoothness condition

$$|\partial_{\mathbf{x}}^\beta V(\mathbf{x})| \le C_\beta |\mathbf{x} - \mathbf{A}_k|^{-1-|\beta|}, \quad \text{for } \mathbf{x} \neq \mathbf{A}_k, \quad \text{and all } |\beta| \geq 0, \tag{2.12}$$

in a neighbourhood of a nucleus.

According to Theorem 1 it is save to assume that $\phi|_{\mathbb{R}^3 \setminus B_{\tilde{R}}} \in \mathcal{S}(\mathbb{R}^3)|_{\mathbb{R}^3 \setminus B_{\tilde{R}}}$, where $B_{\tilde{R}}$ denotes an open ball of radius $\tilde{R} > \max\{|\mathbf{A}_k|\}_{k=1,K}$ centered at the origin. Therefore it is reasonable to restrict the Galerkin discretization to a ball B_R of radius R which contains $B_{\tilde{R}}$ and is sufficiently large such that contributions from $\mathbb{R}^3 \setminus B_R$ can be neglected. On each wavelet level $l \in \mathbb{N}$ we define the set

$$\mathbb{B}_l := \{\mathbf{b} \in \mathbb{Z}^3 : 2^{-l}\mathbf{b} \in B_R\},$$

which specifies the grid points inside the ball B_R. For notational convenience let us assume $R = \mathcal{O}(1)$, i.e. on the coarsest wavelet level $\#\mathbb{B}_0 = \mathcal{O}(1)$.

2.2 s^*-compressibility versus best N-term approximation

Within this section we want to summarize our results concerning the s^*-compressibility of the Galerkin discretization of the Fock operator (1.5) and discuss the relation to previous work on best N-term approximation of solutions of the Hartree-Fock equations. The Fock operator represents a bounded operator from $H^1(\mathbb{R}^3) \to H^{-1}(\mathbb{R}^3)$. By a simple rescaling, the wavelet basis can be converted into a Riesz basis for $H^1(\mathbb{R}^3)$ and the corresponding Galerkin matrix of the Fock operator becomes

$$h_{(j,\mathbf{a}),(l,\mathbf{b})} := 2^{-(j+l)} \int_{\mathbb{R}^3} \gamma_{j,\mathbf{a}}(\mathbf{x})\big(\mathfrak{h}\gamma_{l,\mathbf{b}}\big)(\mathbf{x})d^3x,$$

where $\{h_{(j,\mathbf{a}),(l,\mathbf{b})}\}_{(j,\mathbf{a}),(l,\mathbf{b})\in\Lambda}$ is a bounded operator from $\ell_2(\Lambda) \to \ell_2(\Lambda)$.

Theorem 2. *The infinite symmetric Galerkin matrix of the Fock operator (1.5) in a wavelet basis* $\{\gamma_{j,\mathbf{a}}\}_{(j,\mathbf{a})\in\Lambda}$ *with* $\gamma_{j,\mathbf{a}} \in W_\infty^s(\mathbb{R}^3)$ *and* $r \geq s+1$ *vanishing moments, which furthermore satisfies (2.2) and (2.4), can be split into three parts with different* s^**-compressibility:*

(*i*) *Under the additional assumption of a piecewise smooth wavelet basis, the kinetic energy term* $-\frac{1}{2}\Delta$ *is* s^**-compressible with* $s^* = \min\{\frac{r+1}{3}, \frac{\nu-1}{2}\}$*, where* ν *denotes the Sobolev regularity, i.e.* $\nu = \sup_t\{\gamma_{j,\mathbf{a}} \in H^t(\mathbb{R}^3)\}$*, of the wavelets.*

(*ii*) *The combined Coulomb potential (2.11), which satisfies the asymptotic smoothness condition (2.12), is* s^**-compressible with* $s^* = (s+1)/3$*.*

(*iii*) *The nonlocal exchange operator (1.7) is* s^**-compressible with* $s^* = s/3$ *provided that the functions* $\{\phi_i\}_{i=1,\dots,N/2}$ *are asymptotically smooth in a neighbourhood of a nucleus.*

For an appropriately chosen wavelet basis, the Galerkin matrix of the Fock operator is therefore s^**-compressible with* $s^* = s/3$*.*

The s^*-compressibility of the Laplacian (i) has been already demonstrated by Stevenson [30]. As an additional assumption he required a piecewise smooth wavelet basis which turned out to be necessary in order to achieve the optimal s^*-compressibility. In the following we restrict ourselves to a discussion of the Coulomb potential (2.11) and exchange operator (1.7) where piecewise smooth wavelets will not be assumed.

In order to obtain adaptive algorithms with optimal computational efficiency it is necessary to provide an s^*-compressibility which is greater than the best N-term approximation rates of solutions of the Hartree-Fock equations in the corresponding wavelet basis. The concept of best N-term approximation belongs to the realm of nonlinear approximation theory, cf. Ref. [10] for further details. Loosely speaking we have to consider the best possible approximation of a function $f \in H^1(\Omega)$, for a bounded domain Ω, in the nonlinear subset Σ_N which consists of all possible linear combinations of at most N wavelets[1], i.e.

$$\Sigma_N := \Big\{ \sum_{(j,\mathbf{a})\in\Delta} c_{j,\mathbf{a}}\, \gamma_{j,\mathbf{a}} : \Delta \subset \Lambda, \#\Delta \le N,\, c_{j,\mathbf{a}} \in \mathbb{R} \Big\}.$$

Here, the approximation error

$$\sigma_N(f) := \inf_{f_N \in \Sigma_N} \|f - f_N\|_{H^1}$$

is given with respect to the norm of the Sobolev space $H^1(\Omega)$. Best N-term approximation spaces $A^\alpha_\tau(H^1)$, with $0 < \alpha$, can be defined according to

$$A^\alpha_\tau(H^1) := \{f \in H^1(\Omega) : |f|_{A^\alpha_\tau(H^1)} < \infty\}$$

with semi-norm

$$|f|_{A^\alpha_\tau(H^1)} := \Big(\sum_{N\in\mathbb{N}} \Big(N^\alpha \sigma_N(f) \Big)^\tau N^{-1} \Big)^{\frac{1}{\tau}}.$$

It follows from the definition of the semi-norm that a convergence rate $\sigma_N(f) \sim N^{-\alpha}$ with respect to the number of basis functions can be achieved. In fact this is the best error bound which can be obtained with N degrees of freedom. The spaces $A^\alpha_q(H^1)$ have been identified with standard Besov spaces, i.e. $A^\alpha_q(H^1) \sim B^{3\alpha+1}_\tau\big(L_\tau(\Omega)\big)$ with $\frac{1}{\tau} = \frac{1}{2} + \alpha$. An appropriate wavelet basis for the spaces $A^\alpha_q(H^1)$ must be sufficiently smooth and of order d such that $\alpha < (d-1)/3$, cf. Ref. [10].

Our next theorem is an immediate consequence of Corollary 1 from Ref. [14] and Theorem 1.

Theorem 3. *The solutions* $\{\phi_i\}_{i=1,\dots,N/2}$ *of the Hartree-Fock equations from Theorem 1 belong to the best* N*-term approximation spaces* $\mathcal{A}^\alpha_\tau(H^1)$ *for any* $\alpha > 0$ *and* $\frac{1}{\tau} = \alpha + \frac{1}{2}$*.*

[1] No confusion should arise from our ambiguous notation in this paragraph where N not only denotes the number of electrons but also the number of terms in a best N-term approximation.

Despite the presence of conical singularities at the nuclei, Theorem 3 shows that for an appropriate wavelet basis, the solutions $\{\phi_i\}_{i=1,\ldots,N/2}$ can be approximated with optimal convergence rate.

We are able now to compare the s^*-compressibility of discrete Fock operators with best N-term approximation rates for solutions of the Hartree-Fock equations. For this we restrict ourselves to the common situation of spline wavelets of order $d = s+1$ and with $r \geq s+1$ vanishing moments. In this case we can achieve $\alpha < s/3$ which fits together with $s^* = s/3$ from Theorem 2. Finally we want to mention that due to the nonlocal character of the Hartree-Fock equations, the present considerations are necessary but not sufficient to demonstrate that an adaptive wavelet algorithm with optimal computational complexity exists in the sense of Ref. [8]. To show that such an algorithm actually exists will be the subject of our future research.

2.3 Compression scheme for combined nuclear Coulomb and Hartree potentials

Lemma 2. *Let V be a combined Coulomb potential (2.11) which satisfies the asymptotic smoothness condition (2.12) in a neighbourhood of a nucleus located at $\mathbf{A}_k$. The infinite symmetric Galerkin matrix $\mathcal{V}$ with entries*

$$\mathcal{V}_{(j,\mathbf{a}),(l,\mathbf{b})} := 2^{-(j+l)} \int_{\mathbb{R}^3} \gamma_{j,\mathbf{a}}(\mathbf{x}) V(\mathbf{x}) \gamma_{l,\mathbf{b}}(\mathbf{x}) d^3x, \tag{2.13}$$

is given in a wavelet basis $\{\gamma_{j,\mathbf{a}}\}_{(j,\mathbf{a})\in\Lambda}$ with index set Λ which satisifies (2.2) and (2.4).

For each $q \in \mathbb{N}$ we can compress $\mathcal{V}$ into an infinite symmetric matrix $\mathcal{V}^{(q)}$ where for $l \geq j$ we drop all entries (2.13), and corresponding transposed entries, except of those for which one of the following criteria are satisfied:

(i) $l - j \leq q$.

(ii) $\operatorname{dist}(\operatorname{supp}\gamma_{l,\mathbf{b}}, \mathbf{A}_k) \leq \eta 2^{-l+q}$ *and* $l \leq j_{max}$.

Here $\eta \geq 1$ denotes a scaling parameter which is determined by the size of the mother wavelet. For $s \leq \bar{s} < s+1$ the compression error with respect to the spectral norm can be estimated by

$$\|\mathcal{V} - \mathcal{V}^{(q)}\| \lesssim 2^{-\bar{s}q},$$

and the number of nonzero entries per row and column of $\mathcal{V}^{(q)}$ is bounded on some absolute multiple of

$$j_{max} 2^{3q}, \quad \text{with } j_{max} = \frac{1}{\frac{s+1}{\bar{s}} - 1} q.$$

According to Definition 1, this means in particular that $\mathcal{V}$ is s^*-compressible with $s^* = (s+1)/3$.

An optimal computational complexity can be achieved if we allow for a q dependend $\bar{s}$. The benefit to cost ratio $\bar{s}q/\big(3q + \log_2(j_{max})\big)$ is maximized by taking

$$\alpha_{opt}(q) = \max_{s \leq \bar{s} < s+1} \left\{ \frac{\bar{s}q}{3q + \log_2 q - \log_2\left(\frac{s+1}{\bar{s}} - 1\right)} \right\} \tag{2.14}$$

which is a monotonously increasing function on q with values between

$$\alpha_{opt}(1) = \frac{s}{3 + \log_2 s} \quad (\text{for } s > 1) \quad \text{and} \quad \lim_{q\to\infty} \alpha_{opt}(q) = \frac{s+1}{3}.$$

Table 1: Optimal benefit to cost ratio α_{opt}, cf. (2.14), for various wavelet regularities s and different values of the refinement paramter q, cf. conditions (i) and (ii) in Lemma 2.

	q					
s	1	2	5	10	100	∞
1	0.33	0.36	0.44	0.51	0.63	0.67
3	0.65	0.72	0.88	1.01	1.27	1.33
5	0.94	1.07	1.33	1.52	1.90	2

Proof. For notational convenience we introduce the residuum matrix

$$\Delta\mathcal{V}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})} := \mathcal{V}_{(j,\mathbf{a}),(l,\mathbf{b})} - \mathcal{V}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}$$

with $(j,\mathbf{a})$, $(l,\mathbf{b}) \in \Lambda$. In the following we assume w.l.o.g. the presence of a single nucleus located at the origin. According to our criteria (i) and (ii), we have to consider for $l > j+q$ the case $\operatorname{dist}(\operatorname{supp}\gamma_{j,\mathbf{a}}, \mathbf{0}) > \eta 2^{-j}$ which contributes

$$\begin{aligned}
\sum_{\mathbf{b}\in\mathbb{B}_l} \left|\Delta\mathcal{V}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| &\lesssim 2^{-(l+j)} 2^{-(\frac{3}{2}+s)l} \sum_{\mathbf{b}\in\mathbb{B}_l} \left\|V\gamma_{j,\mathbf{a}}\right\|_{W^s_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} \\
&\lesssim 2^{-(l+j)} 2^{-(\frac{3}{2}+s)l} 2^{3(l-j)} 2^{(\frac{5}{2}+s)j} \\
&\simeq 2^{-(s-\frac{1}{2})l} 2^{(s-\frac{3}{2})j} \\
&\lesssim 2^{-(s-2)q} 2^{\frac{1}{2}j} 2^{-\frac{3}{2}l}
\end{aligned}$$

and the case $\operatorname{dist}(\operatorname{supp}\gamma_{j,\mathbf{a}}, \mathbf{0}) \le \eta 2^{-j}$, $\operatorname{dist}(\operatorname{supp}\gamma_{l,\mathbf{b}}, \mathbf{0}) \ge \eta 2^{-l+q}$, i.e. $\mathbf{b}$ belongs to the index set

$$\Lambda_{j\mathbf{a},l} := \{\mathbf{b} \in \mathbb{B}_l : \operatorname{supp}\gamma_{l,\mathbf{b}} \cap \operatorname{supp}\gamma_{j,\mathbf{a}} \neq \emptyset \;\text{ and } |\mathbf{b}| \ge \eta 2^q\},$$

where we get the estimate

$$\begin{aligned}
\sum_{\mathbf{b}\in\Lambda_{j\mathbf{a},l}} \left|\Delta\mathcal{V}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| &\lesssim 2^{-(l+j)} 2^{-(\frac{3}{2}+s)l} \sum_{\mathbf{b}\in\Lambda_{j\mathbf{a},l}} \left\|V\gamma_{j,\mathbf{a}}\right\|_{W^s_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} \\
&\lesssim 2^{-(l+j)} 2^{-(\frac{3}{2}+s)l} \sum_{\mathbf{b}\in\Lambda_{j\mathbf{a},l}} \sup_{|\alpha_1|+|\alpha_2|\le s} \left\{ 2^{(\frac{3}{2}+|\alpha_1|)j} |2^{-l}\mathbf{b}|^{-1-|\alpha_2|} \right\} \\
&\lesssim 2^{-(l+j)} 2^{-(\frac{3}{2}+s)l} \sup_{|\alpha_1|+|\alpha_2|\le s} \left\{ 2^{(\frac{3}{2}+|\alpha_1|)j} \sum_{\mathbf{b}\in\Lambda_{j\mathbf{a},l}} |2^{-l}\mathbf{b}|^{-1-|\alpha_2|} \right\} \\
&\lesssim 2^{-(\frac{5}{2}+s)l} 2^{-j} 2^{3l} 2^{\frac{3}{2}j} 2^{(s-2)(l-q)} \\
&= 2^{-(s-2)q} 2^{\frac{1}{2}j} 2^{-\frac{3}{2}l}.
\end{aligned} \tag{2.15}$$

In the third line of the estimate we have used

$$\sum_{\mathbf{b}\in\Lambda_{j\mathbf{a},l}} |2^{-l}\mathbf{b}|^{-1-|\alpha|} \lesssim 2^{3l} \int_{\eta 2^{-l+q}}^{\eta 2^{-j}} t^{1-|\alpha|} dt \lesssim 2^{3l} \left\{ \begin{array}{cl} 2^{(|\alpha|-2)j} & \text{for } |\alpha| < 2 \\ l-j-q & \text{for } |\alpha| = 2 \\ 2^{(|\alpha|-2)(l-q)} & \text{for } |\alpha| > 2 \end{array} \right. .$$

It follows from the previous estimates that the contributions of $l > j+q$ to the spectral norm, using

Schur's lemma, are given by

$$\sup_{j\geq 0}\Big\{\sum_{l>j+q} 2^{p(j-l)} \sum_{\mathbf{b}\in\mathbb{B}_l} \big|\Delta\mathcal{V}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\big|\Big\} \lesssim 2^{-(s-2)q}\sup_{j\geq 0}\Big\{\sum_{l>j+q} 2^{-(p+\frac{3}{2})l}2^{(p+\frac{1}{2})j}\Big\}$$
$$\lesssim 2^{-(s-2)q}\sup_{j\geq 0}\Big\{2^{-(p+\frac{3}{2})q}2^{-j}\Big\} \quad \text{for } p > -\tfrac{3}{2}$$
$$\lesssim 2^{-(s+p-\frac{1}{2})q}. \tag{2.16}$$

For $l \leq j-q$, we have to consider the case $\operatorname{dist}(\operatorname{supp}\gamma_{j,\mathbf{a}},\mathbf{0}) \geq \eta 2^{-j+q}$. Because of $\#\{\gamma_{l,\mathbf{b}} : \operatorname{supp}\gamma_{j,\mathbf{a}} \cap \operatorname{supp}\gamma_{l,\mathbf{b}} \neq \emptyset\} = \mathcal{O}(1)$ it is sufficient to estimate on each level l only a single entry

$$\big|\Delta\mathcal{V}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\big| \lesssim 2^{-(l+j)}2^{-(\frac{3}{2}+s)j}\big\|V\gamma_{l,\mathbf{b}}\big\|_{W^s_\infty(\operatorname{supp}\gamma_{j,\mathbf{a}})}$$
$$\lesssim 2^{-(l+j)}2^{-(\frac{3}{2}+s)j}\sup_{|\alpha_1|+|\alpha_2|\leq s}\Big\{2^{(\frac{3}{2}+|\alpha_1|)l}2^{(1+|\alpha_2|)(j-q)}\Big\}$$
$$\lesssim 2^{-(l+j)}2^{-(\frac{3}{2}+s)j}2^{\frac{3}{2}l}2^{(1+s)(j-q)}$$
$$\lesssim 2^{-(s+1)q}2^{-\frac{3}{2}j}2^{\frac{1}{2}l}. \tag{2.17}$$

Alltogether the contribution of $l < j-q$ to the spectral norm is

$$\sup_{j\geq 0}\Big\{\sum_{0\leq l<j-q} 2^{p(j-l)} \sum_{\mathbf{b}\in\mathbb{B}_l} \big|\Delta\mathcal{V}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\big|\Big\} \lesssim 2^{-(s+1)q}\sup_{j\geq 0}\Big\{\sum_{0\leq l<j-q} 2^{-(p-\frac{1}{2})l}2^{(p-\frac{3}{2})j}\Big\}$$
$$\lesssim \sup_{j\geq 0}\begin{cases} 2^{-(s+1)q}2^{(p-\frac{3}{2})j} & \text{for } \frac{1}{2} < p \leq \frac{3}{2} \\ (j-q)2^{-(s-p+\frac{3}{2})q}2^{-j} & \text{for } p \leq \frac{1}{2} \end{cases}$$
$$\lesssim \begin{cases} 2^{-(s+1)q} & \text{for } \frac{1}{2} < p \leq \frac{3}{2} \\ 2^{-(s-p+\frac{3}{2})q} & \text{for } p \leq \frac{1}{2} \end{cases}. \tag{2.18}$$

The adaptive refinement cannot be maintained up to arbitrarily fine levels while keeping in each row and column the number of nonzero entries bounded on some absolute multiple of $q2^{3q}$, see e.g. Fig. 1 a). Therefore it is necessary to introduce a finest wavelet level j_{max} beyond which adaptive refinements are not taken into account. A comparison of (2.16) and (2.18) shows that we can restrict the following considerations to $\frac{1}{2} < p \leq \frac{3}{2}$. For columns of $\mathcal{V}$ with index $j > j_{max}$, we can easily estimate the contribution to the spectral norm using (2.15) and (2.17) with $q = 0$ and the following estimate

$$\big|\Delta\mathcal{V}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\big| \lesssim \begin{cases} 2^{\frac{1}{2}j}2^{-\frac{3}{2}l} & \text{for } l > j+q \\ 2^{-\frac{3}{2}j}2^{\frac{1}{2}l} & \text{for } l < j-q \end{cases},$$

where we assume $\#\{\mathbf{b}\in\mathbb{B}_l : \operatorname{dist}(\operatorname{supp}\gamma_{l,\mathbf{b}},\mathbf{0}) \leq \eta 2^{-l}\} = \mathcal{O}(1)$. This yields the contributions

$$\sup_{j>j_{max}}\Big\{\sum_{l>j+q} 2^{p(j-l)} \sum_{\mathbf{b}\in\mathbb{B}_l} \big|\Delta\mathcal{V}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\big|\Big\} \lesssim \sup_{j>j_{max}}\sum_{l>j+q} 2^{p(j-l)}2^{\frac{1}{2}j}2^{-\frac{3}{2}l}$$
$$\lesssim 2^{-(p+\frac{3}{2})q}2^{-j_{max}} \quad \text{for } p > -\tfrac{3}{2} \tag{2.19}$$

and

$$\sup_{j>j_{max}}\Big\{\sum_{0\leq l<j-q} 2^{p(j-l)} \sum_{\mathbf{b}\in\mathbb{B}_l} \big|\Delta\mathcal{V}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\big|\Big\} \lesssim \sup_{j>j_{max}}\sum_{0\leq l<j-q} 2^{p(j-l)}2^{-\frac{3}{2}j}2^{\frac{1}{2}l}$$
$$\lesssim 2^{-(\frac{3}{2}-p)j_{max}} \quad \text{for } \tfrac{1}{2} < p \leq \tfrac{3}{2}, \tag{2.20}$$

where it can be seen that the estimate for $l < j-q$ dominates. It is convenient to define a parameter $\epsilon := \frac{3}{2} - p$ such that $0 < \epsilon < 1$. In order to get the number of nonzero entries in each row and column bounded such that an overall accuracy

$$\|\Delta\mathcal{V}\| \lesssim 2^{-(s+1-\epsilon)q}, \quad \text{for } 0 < \epsilon < 1 \tag{2.21}$$

is achieved, we must require

$$\big(s+1-\epsilon\big)q = \epsilon j_{max}.$$

This means that the number of nonzero entries in each row and column is bounded on some absolute multiple of

$$\frac{s+1-\epsilon}{\epsilon}\, q2^{3q}. \tag{2.22}$$

It is easy to see that up to a constant (2.21) and (2.22) also applies for $\epsilon = 1$. Altogether, this demonstrates the s^*-compressibility of $\mathcal{V}$ for $s^* = (s+1)/3$. Because for every $0 < \alpha < (s+1)/3$, cf. Definition 1, we can find an $\bar{s} := s+1-\epsilon \geq s$, with $\alpha < \bar{s}/3 < (s+1)/3$, such that the requirements of Definition 1 are satisfied.

Obviously, the number of nonzero entries in each row and column increases beyond all bounds for $\alpha \to (s+1)/3$, i.e. $\bar{s} \to s+1$. In order to study the relation between (2.21) and (2.22) more closely we consider the benefit to cost ratio

$$R_q(\bar{s}) = \left\{ \frac{\bar{s}q}{3q + \log_2 q - \log_2\left(\frac{s+1}{\bar{s}} - 1\right)} \right\}, \quad \text{for } s \leq \bar{s} < s+1,$$

depending on $q \in \mathbb{N} \setminus \{0\}$. For each $q \geq 1$ the benefit to cost ratio has a unique maximum $\alpha_{opt}(q) := \max R_q$. Depending on the wavelet regularity s there exists a $q_0 \in \mathbb{N}$ such that for $1 \leq q < q_0$ the maximum is at $\bar{s} = s$ with

$$\alpha_{opt}(q) = \frac{sq}{3q + \log_2 q + \log_2 s}$$

and for $q \geq q_0$ it is the unique solution of

$$\ln \frac{1}{\frac{s+1}{\bar{s}} - 1} = \frac{1}{\frac{s+1}{\bar{s}} - 1} - 3(\ln 2)q - \ln q + 1$$

with $s \leq \bar{s} < s+1$. The optimal benefit to cost ratio α_{opt} is monotonously increasing with q and finally approaches

$$\lim_{q\to\infty} \alpha_{opt}(q) = \frac{s+1}{3}.$$

In Table 2.3 we have listed α_{opt} for various values of q and different regularities s. □

2.4 Compression scheme for the nonlocal exchange operator

In order to discuss the s^*-compressibility of the exchange operator (1.7) it is an essential requirement that the eigenfunctions ϕ_i, $i = 1, \dots N/2$, which belong to the $N/2$ lowest eigenvalues of the Fock operator are asymptotically smooth away from the nuclei, i.e., $|\partial^\beta_{\mathbf{x}}\phi(\mathbf{x})| \leq C_\beta|\mathbf{x}-\mathbf{A}_k|^{1-|\beta|}$ for $\mathbf{x} \neq \mathbf{A}_k$, $k = 1, \dots, K$ and $|\beta| \geq 1$. The eigenfunctions introduce nondiagonal singularities into the kernel which compel rather tedious estimates for different arrangement of wavelets in the neighbourhoods of nuclei. Various possible arrangements are shown in Figure 2. To treat adaptivity as flexible as possible we have introduced three additional distance parameters $0 \leq q', \tilde{q}, \tilde{q}' \leq q$ into our adaptive

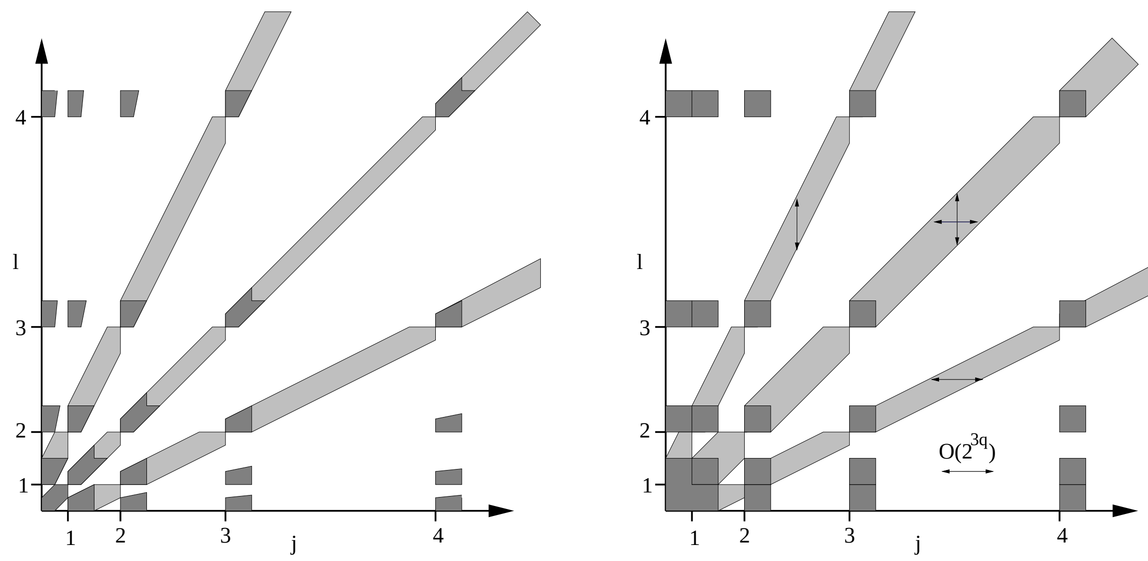

Figure 1: Nonzero entries of compressed parts of the Fock matrix. Galerkin discretization of combined Coulomb and Hartree potential (a), and exchange operator (b). Light shaded regions correspond to $|l - j| \leq q$ with $q = 1$ and dark shaded regions highlight adative refinements near nuclei.

estimates. For further reference we define minimal distances $d_{j\mathbf{a},\star}$ between supports of wavelets and the nuclei, i.e.

$$d_{j\mathbf{a},\star} := \operatorname{dist}(\operatorname{supp}\gamma_{j,\mathbf{a}}, \{\mathbf{A}_k\}_{k=1,K})$$

and distances $d_{j\mathbf{a},l\mathbf{b}}$ between the supports of two wavelets, i.e.

$$d_{j\mathbf{a},l\mathbf{b}} := \operatorname{dist}(\operatorname{supp}\gamma_{j,\mathbf{a}}, \operatorname{supp}\gamma_{l,\mathbf{b}}).$$

Our main result for the s^*-compressibility of the exchange operator can be summerized in the following lemma.

Lemma 3. *The infinite symmetric Galerkin matrix of the exchange operator (1.7) with entries*

$$\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})} := 2^{-(j+l)} \int\limits_{\mathbb{R}^3} \gamma_{j,\mathbf{a}}(\mathbf{x})\big(U\gamma_{l,\mathbf{b}}\big)(\mathbf{x})d^3x,$$

in a wavelet basis $\{\gamma_{j,\mathbf{a}}\}_{(j,\mathbf{a})\in\Lambda}$ with $r \geq s+1$ vanishing moments, which furthermore satisfies (2.2) and (2.4), is s^*-compressible with $s^* = s/3$.

This follows from the fact that for each $q \in \mathbb{N}$ we can compress $\mathcal{U}$ into a symmetric matrix $\mathcal{U}^{(q)}$ with spectral norm estimate

$$\|\mathcal{U} - \mathcal{U}^{(q)}\| \lesssim 2^{-(s-\epsilon)q} \quad \textit{for any } \epsilon > 0,$$

where we drop all entries except of those with indices $j, l \leq j_{max}$ for which one the following distance criteria applies

- *$j + q < l$, cf. case D in Fig. 2.*

  (i) $d_{l\mathbf{b},\star} \leq \eta 2^{-l+q}$ *and* $d_{j\mathbf{a},\star} \leq \eta 2^{-j+q}$

- *$j \leq l \leq j + q$, cf. case C (a), D and E in Fig. 2.*

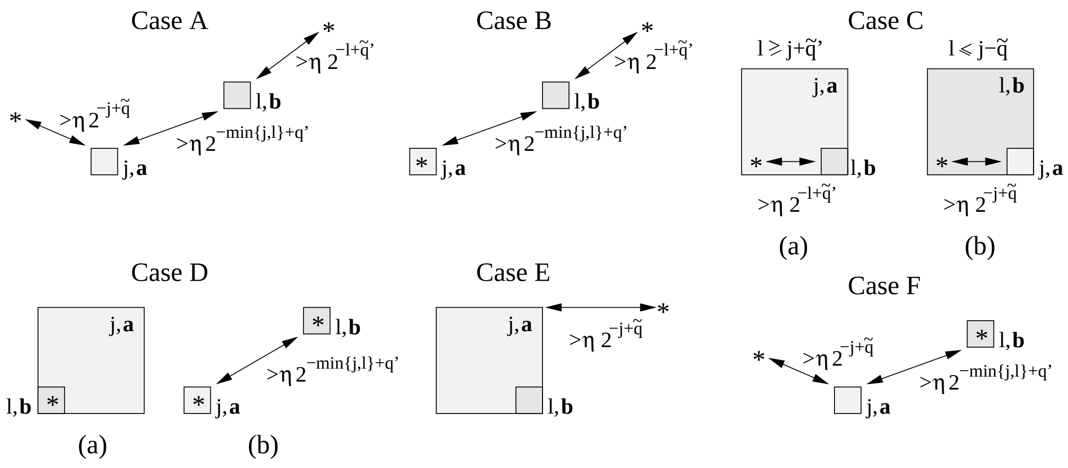

Figure 2: Possible arrangements of wavelets and nuclei which have been taken care of in the proof of Lemma 3.

(i) $d_{j\mathbf{a},l\mathbf{b}} \leq \eta 2^{-l+q}$

(ii) $d_{l\mathbf{b},\star} \leq \eta 2^{-l+q}$ *and* $d_{j\mathbf{a},\star} \leq \eta 2^{-j+q}$

- $j - q \leq l < j$, *cf. case C (b), D and E in Fig. 2.*

  (i) $d_{j\mathbf{a},l\mathbf{b}} \leq \eta 2^{-j+q}$

  (ii) $d_{j\mathbf{a},\star} \leq \eta 2^{-j+q}$ *and* $d_{l\mathbf{b},\star} \leq \eta 2^{-l+q}$

- $0 \leq l < j - q$, *cf. case D in Fig. 2.*

  (i) $d_{j\mathbf{a},\star} \leq \eta 2^{-j+q}$ *and* $d_{l\mathbf{b},\star} \leq \eta 2^{-l+q}$

The finest level $j_{max} = c_\alpha q$*, with appropriate constant for any compression rate* α *with* $0 < \alpha < s^*$*, can be chosen such that* $\mathcal{U}^{(q)}$ *has in each row and column at most* $O(q2^{3q})$ *nonzero entries.*

In the following we assume that $\mathbf{a} \notin \mathbb{B}_j$ if $d_{j\mathbf{a},\star} \geq \eta$, where $\eta \geq 1$ denotes a scaling parameter which is determined by the size of the mother wavelet. To simplify our notation, we consider in the following w.l.o.g. the Galerkin discretization of the reduced exchange kernel

$$\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})} := 2^{-(j+l)} \int\limits_{\mathbb{R}^3} \int\limits_{\mathbb{R}^3} \gamma_{j,\mathbf{a}}(\mathbf{x}) \frac{\phi(\mathbf{x})\phi(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \gamma_{l,\mathbf{b}}(\mathbf{y}) d^3x d^3y \tag{2.23}$$

where the function ϕ is asymptotically smooth away from the nuclei.

2.5 Estimates for partial sums over rows of the exchange matrix

As a first step of the proof of Lemma 3 we estimate, for fixed row index $(j, \mathbf{a})$ and fixed column level l, the partial sums $\sum_{\mathbf{b}\in\Lambda_{j\mathbf{a},l}} |\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}|$ where index sets $\Lambda_{j\mathbf{a},l}$ are specified according to the different arrangements depicted in Fig. 2.

Case A)

For columns with $d_{j\mathbf{a},\star} \geq \eta 2^{-j+\tilde{q}}$ we consider the set

$$\Lambda^{(a)}_{j\mathbf{a},l} := \left\{ \mathbf{b} \in \mathbb{B}_l : d_{j\mathbf{a},l\mathbf{b}} \geq \eta 2^{-\min\{j,l\}+q'} \text{ and } d_{l\mathbf{b},\star} \geq \eta 2^{-l+\tilde{q}'} \right\}, \tag{2.24}$$

with $\mathbb{N} \ni q', \tilde{q}, \tilde{q}' \leq q$. In the following let us denote with $\square_{j,\mathbf{a}} := \operatorname{supp} \gamma_{j,\mathbf{a}}$ and $\square_{l,\mathbf{b}} := \operatorname{supp} \gamma_{l,\mathbf{b}}$ the supports of the wavelets. The supports do not overlap, therefore Proposition 1 (ii) can be applied with respect to both variables

$$\sum_{\mathbf{b} \in \Lambda_{j\mathbf{a},l}^{(a)}} \left| \mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})} \right| \lesssim 2^{-(r+\frac{5}{2})(j+l)} \sum_{\mathbf{b} \in \Lambda_{j\mathbf{a},l}^{(a)}} \left\| \partial_x^r \partial_y^r \frac{\phi(\mathbf{x})\phi(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \right\|_{L_\infty(\square_{j,\mathbf{a}} \times \square_{l,\mathbf{b}})}.$$

The sum can be estimated according to

$$\begin{aligned} \sum_{\mathbf{b} \in \Lambda_{j\mathbf{a},l}^{(a)}} \left\| \partial_x^r \partial_y^r \frac{\phi(\mathbf{x})\phi(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \right\|_{L_\infty(\square_{j,\mathbf{a}} \times \square_{l,\mathbf{b}})} &\lesssim \sum_{\mathbf{b} \in \Lambda_{j\mathbf{a},l}^{(a)}} \sup_{\substack{s_1+s_2=r \\ t_1+t_2=r}} \Big\{ \max\big\{2^{(s_1-1)(j-\tilde{q})}, 1\big\} \\ &\quad \times \max\big\{2^{(t_1-1)(l-\tilde{q}')}, 1\big\} \left\| |\mathbf{x}-\mathbf{y}|^{-1-s_2-t_2} \right\|_{L_\infty(\square_{j,\mathbf{a}} \times \square_{l,\mathbf{b}})} \Big\} \\ &\lesssim \sup_{\substack{s_1+s_2=r \\ t_1+t_2=r}} \Big\{ \max\big\{2^{(s_1-1)(j-\tilde{q})}, 1\big\} \max\big\{2^{(t_1-1)(l-\tilde{q}')}, 1\big\} \\ &\quad \times 2^{3l} \int_{\eta 2^{-\min\{j,l\}+q'}}^{R} u^{1-s_2-t_2} du \Big\} \\ &\lesssim 2^{3l} \sup_{\substack{s_1+s_2=r \\ t_1+t_2=r}} \Big\{ \max\big\{2^{(s_1-1)(j-\tilde{q})}, 1\big\} \max\big\{2^{(t_1-1)(l-\tilde{q}')}, 1\big\} \\ &\quad \times \max\big\{2^{(s_2+t_2-2)(\min\{j,l\}-q')}, 1\big\} \Big\}. \end{aligned}$$

A further simplification is achieved using

$$\max\big\{2^{(s_2+t_2-2)(\min\{j,l\}-q')}, 1\big\} \lesssim \max\big\{2^{(s_2-1)(\min\{j,l\}-q')}, 1\big\} \max\big\{2^{(t_2-1)(\min\{j,l\}-q')}, 1\big\}$$

to get

$$\begin{aligned} \sum_{\mathbf{b} \in \Lambda_{j\mathbf{a},l}^{(a)}} \left| \mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})} \right| &\lesssim 2^{-(r+\frac{5}{2})j} 2^{-(r-\frac{1}{2})l} \sup_{\substack{s_1+s_2=r \\ t_1+t_2=r}} \Big\{ \max\big\{2^{(s_1-1)(j-\tilde{q})}, 1\big\} \max\big\{2^{(t_1-1)(l-\tilde{q}')}, 1\big\} \\ &\quad \times \max\big\{2^{(s_2-1)(\min\{j,l\}-q')}, 1\big\} \max\big\{2^{(t_2-1)(\min\{j,l\}-q')}, 1\big\} \Big\}. \end{aligned} \tag{2.25}$$

From this estimate we can easily obtain some special cases which are required in the following.

(a1) For $l > j + q$ we have $q' = 0$ in order to keep the number of nonzero entries per row and column within the tolerated limit. Using $l - q > j \geq 0$, we get from (2.25)

$$\begin{aligned} \sum_{\mathbf{b} \in \Lambda_{j\mathbf{a},l}^{(a)}} \left| \mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})} \right| &\lesssim 2^{-(r+\frac{5}{2})j} 2^{-(r-\frac{1}{2})l} 2^{(r-1)j} 2^{(r-1)(l-\tilde{q}')} \\ &\simeq 2^{-(r-1)\tilde{q}'} 2^{-\frac{7}{2}j} 2^{-\frac{1}{2}l}. \end{aligned} \tag{2.26}$$

(a2) For $j \leq l \leq j + q$, adaptive refinement along the diagonal with $q' = q - l + j$ is possible. We can restrict ourselves to the case $\tilde{q} = 0$ and obtain from (2.25)

$$\begin{aligned} \sum_{\mathbf{b} \in \Lambda_{j\mathbf{a},l}^{(a)}} \left| \mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})} \right| &\lesssim 2^{-(r+\frac{5}{2})j} 2^{-(r-\frac{1}{2})l} 2^{(r-1)(l-\tilde{q}')} 2^{(r-1)j} \\ &\simeq 2^{-(r-1)\tilde{q}'} 2^{-\frac{7}{2}j} 2^{-\frac{1}{2}l} \end{aligned} \tag{2.27}$$

(a3) For $j-q \le l < j$, adaptive refinement with $q' = q-j+l$ can be chosen. In the following (2.25) restricts to the case $\tilde{q}' = 0$ where we obtain

$$\sum_{\mathbf{b}\in\Lambda^{(a)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-(r+\frac{5}{2})j}2^{-(r-\frac{1}{2})l}2^{(r-1)(j-\tilde{q})}2^{(r-1)l} \simeq 2^{-(r-1)\tilde{q}}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.28}$$

(a4) For $0 \le l < j-q$, we have $q' = 0$ and (2.25) yields

$$\sum_{\mathbf{b}\in\Lambda^{(a)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-(r-1)\tilde{q}}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.29}$$

Case B)

For columns with $d_{j\mathbf{a},\star} \le \eta 2^{-j+\tilde{q}}$ we consider the set

$$\Lambda^{(b)}_{j\mathbf{a},l} := \left\{\mathbf{b}\in\mathbb{B}_l : d_{j\mathbf{a},l\mathbf{b}} \ge \eta 2^{-\min\{j,l\}+q'} \text{ and } d_{l\mathbf{b},\star} \ge \eta 2^{-l+\tilde{q}'}\right\}. \tag{2.30}$$

Using $\phi\in C^{0,1}(\mathbb{R}^3)$ and Proposition 1 we obtain the estimate

$$\sum_{\mathbf{b}\in\Lambda^{(b)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-\frac{7}{2}j}2^{-(r+\frac{5}{2})l} \sum_{\mathbf{b}\in\Lambda^{(b)}_{j\mathbf{a},l}} \left\|\partial_x^1\partial_y^r \frac{\phi(\mathbf{x})\phi(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\right\|_{L_\infty(\square_{j,\mathbf{a}}\times\square_{l,\mathbf{b}})}.$$

The sum can be further estimated

$$\begin{aligned}
&\sum_{\mathbf{b}\in\Lambda^{(b)}_{j\mathbf{a},l}} \left\|\partial_x^1\partial_y^r \frac{\phi(\mathbf{x})\phi(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\right\|_{L_\infty(\square_{j,\mathbf{a}}\times\square_{l,\mathbf{b}})}\\
&\quad\lesssim \sum_{\mathbf{b}\in\Lambda^{(b)}_{j\mathbf{a},l}} \sup_{s_1+s_2=r}\left\{\max\{2^{(s_1-1)(l-\tilde{q}')},1\}\left\||\mathbf{x}-\mathbf{y}|^{-2-s_2}\right\|_{L_\infty(\square_{j,\mathbf{a}}\times\square_{l,\mathbf{b}})}\right\}\\
&\quad\lesssim \sup_{s_1+s_2=r}\left\{\max\{2^{(s_1-1)(l-\tilde{q}')},1\}2^{3l}\int_{\eta 2^{-\min\{j,l\}+q'}}^{R} u^{-s_2}\,du\right\}\\
&\quad\lesssim 2^{3l}\sup_{s_1+s_2=r}\left\{\max\{2^{(s_1-1)(l-\tilde{q}')},1\}\max\{2^{(s_2-1)(\min\{j,l\}-q')},1\}\right\},
\end{aligned}$$

which together with the prefactor yields the general estimate

$$\sum_{\mathbf{b}\in\Lambda^{(b)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-\frac{7}{2}j}2^{-(r-\frac{1}{2})l}\sup_{s_1+s_2=r}\left\{\max\{2^{(s_1-1)(l-\tilde{q}')},1\}\max\{2^{(s_2-1)(\min\{j,l\}-q')},1\}\right\}. \tag{2.31}$$

We can now specify some special cases which are required in the following.

(b1) For $l > j+q$ we have to take $q' = 0$ in (2.31) to get

$$\sum_{\mathbf{b}\in\Lambda^{(b)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-(r-1)\tilde{q}'}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.32}$$

(b2) For $j \le l \le j+q$ adaptive refinement requires $q' = q - l + j$ and (2.31) yields

$$\sum_{\mathbf{b}\in\Lambda^{(b)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-(r-1)\tilde{q}'} 2^{-\frac{7}{2}j} 2^{-\frac{1}{2}l}. \tag{2.33}$$

(b3) For $0 \le l < j$ we can restrict ourselves in (2.31) to the case $q' = \tilde{q}' = q$ and obtain

$$\sum_{\mathbf{b}\in\Lambda^{(b)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-(r-1)q} 2^{-\frac{7}{2}j} 2^{-\frac{1}{2}l}. \tag{2.34}$$

Case C)
This special case only occurs for $l > j + \tilde{q}'$ or $l < j - \tilde{q}$. In the first case with $l > j + \tilde{q}'$ we take into account columns with $d_{j\mathbf{a},\star} \le \eta 2^{-j}$ and the corresponding sets

$$\Lambda^{(c)}_{j\mathbf{a},l} := \left\{\mathbf{b} \in \mathbb{B}_l : d_{j\mathbf{a},l\mathbf{b}} \le \eta 2^{-j} \text{ and } d_{l\mathbf{b},\star} > \eta 2^{-l+\tilde{q}'}\right\}. \tag{2.35}$$

In order to estimate matrix elements we consider around each nucleus $k \in \{1,\dots,K\}$ with $\operatorname{dist}\{\mathbf{A}_k, \operatorname{supp}\gamma_{j,\mathbf{a}}\} < \eta 2^{-j}$ a ball $B_r(k)$ of radius $r = \eta 2^{-l+q''}$ with $q'' \le \tilde{q}'$. The union of these balls form a Lipschitz domain Ω_r with boundary $\partial\Omega_r$. Next we consider a ball B_R with center in $\operatorname{supp}\gamma_{l,\mathbf{b}}$ and radius $R = 5\eta 2^{-j}$. We denote the boundary of the ball by $\partial\Omega_R$ and consider now the Lipschitz domain $\Omega := B_R \setminus \overline{\Omega}_r$ with boundary $\partial\Omega = \partial\Omega_R \cup \partial\Omega_r$. By construction we have

$$\operatorname{dist}\{\partial\Omega_r, \operatorname{supp}\gamma_{l,\mathbf{b}}\} \ge \eta 2^{-l+\tilde{q}'} \quad \text{and} \quad \operatorname{dist}\{\partial\Omega_R, \operatorname{supp}\gamma_{l,\mathbf{b}}\} \ge \eta 2^{-j}.$$

Furthemore we can assume $\operatorname{vol}\Omega_r \sim 2^{-3(l-q'')}$ and $\operatorname{area}\partial\Omega_r \sim 2^{-2(l-q'')}$. Let us first consider the domain Ω_r where we obtain the estimate

$$\begin{aligned}
&\left|\int_{\mathbb{R}^3}\int_{\Omega_r} \gamma_{j,\mathbf{a}}(\mathbf{x}) \frac{\phi(\mathbf{x})\phi(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\gamma_{l,\mathbf{b}}(\mathbf{y}) d^3x d^3y\right| \\
&\qquad \lesssim 2^{-(r+\frac{3}{2})l} \left\|\partial_y^r\left(\phi(\mathbf{y})\int_{\Omega_r} \frac{\phi(\mathbf{x})\gamma_{j,\mathbf{a}}(\mathbf{x})}{|\mathbf{x}-\mathbf{y}|}d^3x\right)\right\|_{L_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} \\
&\qquad \lesssim 2^{-(r+\frac{3}{2})l} \sup_{t_1+t_2=r} \left\| \left|\partial_y^{t_1}\phi(\mathbf{y})\right| \, 2^{\frac{3}{2}j} \int_{\Omega_r} |\mathbf{x}-\mathbf{y}|^{-1-t_2} d^3x \right\|_{L_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} \\
&\qquad \lesssim 2^{-(r+\frac{3}{2})l} \sup_{t_1+t_2=r} \left\{ \max\{2^{(t_1-1)(l-\tilde{q}')}, 1\} 2^{\frac{3}{2}j} 2^{-3(l-q'')} 2^{(t_2+1)(l-\tilde{q}')} \right\} \\
&\qquad \lesssim 2^{-(r+1)\tilde{q}'} 2^{3q''} 2^{\frac{3}{2}j} 2^{-\frac{7}{2}l}.
\end{aligned} \tag{2.36}$$

The enveloping domain Ω yields the estimate

$$\begin{aligned}
&\left|\int_{\mathbb{R}^3}\int_{\Omega} \gamma_{j,\mathbf{a}}(\mathbf{x}) \frac{\phi(\mathbf{x})\phi(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\gamma_{l,\mathbf{b}}(\mathbf{y}) d^3x d^3y\right| \\
&\qquad \lesssim 2^{-(s+\frac{5}{2})l} \left\|\partial_y^{s+1}\left(\phi(\mathbf{y})\int_{\Omega} \frac{\phi(\mathbf{x})\gamma_{j,\mathbf{a}}(\mathbf{x})}{|\mathbf{x}-\mathbf{y}|}d^3x\right)\right\|_{L_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} \\
&\qquad \lesssim 2^{-(s+\frac{5}{2})l} \max\left\{ 2^{s(l-\tilde{q}')} 2^{-\frac{1}{2}j}, \sup_{t_1+t_2=s} \left(\max\{2^{(t_1-1)(l-q'')}, 1\} 2^{(t_2+\frac{1}{2})j} \right) \right\} \\
&\qquad \lesssim 2^{-(s+\frac{5}{2})l} \max\{2^{s(l-\tilde{q}')} 2^{-\frac{1}{2}j}, 2^{(s-1)(l-q'')} 2^{\frac{1}{2}j}, 2^{(\frac{1}{2}+s)j}\} \\
&\qquad \lesssim 2^{\frac{1}{2}j} 2^{-(s+\frac{5}{2})l} \max\{2^{s(l-\tilde{q}')}, 2^{(s-1)(l-q'')}\},
\end{aligned} \tag{2.37}$$

where we have used

$$\left\| \int_\Omega \frac{\phi(\mathbf{x})\gamma_{j,\mathbf{a}}(\mathbf{x})}{|\mathbf{x}-\mathbf{y}|} d^3x \right\|_{L_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} \lesssim 2^{\frac{3}{2}j} \int_0^{2^{-j}} u\,du \lesssim 2^{-\frac{1}{2}j}, \tag{2.38}$$

and for $0 < t \le s+1$

$$\begin{aligned}\left\| \partial_y^t \int_\Omega \frac{\phi(\mathbf{x})\gamma_{j,\mathbf{a}}(\mathbf{x})}{|\mathbf{x}-\mathbf{y}|} d^3x \right\|_{L_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} &\lesssim 2^{-j} \left\| \partial_i^{t-1}(\phi\gamma_{j,\mathbf{a}}) \right\|_{L_\infty(\Omega)} + \delta_{t,s+1} \left\| \partial_i^{s-1}(\phi\gamma_{j,\mathbf{a}}) \right\|_{L_\infty(\Omega)} \\ &\lesssim 2^{-j} \sup_{t_1+t_2=t-1} \left\{ \max\left\{2^{(t_1-1)(l-q'')}, 1\right\} 2^{(t_2+\frac{3}{2})j} \right\}.\end{aligned}$$

The latter estimate is a consequence of Proposition 3 and the following volume and surface estimates

$$\left\| \int_\Omega \frac{1}{|\mathbf{x}-\mathbf{y}|^2} d^3y \right\|_{L_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} \lesssim 2^{-j},$$

$$\left\| \int_{\partial\Omega_r} \frac{1}{|\mathbf{x}-\mathbf{y}|^2} ds \right\|_{L_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} \lesssim 2^{2(l-\tilde{q}')} \operatorname{area} \partial\Omega_r \lesssim 2^{2(l-\tilde{q}')} 2^{-2(l-q'')} \lesssim 1,$$

$$\left\| \int_{\partial\Omega_R} \frac{1}{|\mathbf{x}-\mathbf{y}|^2} ds \right\|_{L_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} \lesssim 2^{2j} \operatorname{area} \partial\Omega_R \lesssim 1.$$

It remains to determine the parameter q'' in order to obtain balanced estimates (2.36) and (2.37). A comparison of the two estimates reveals that (2.37) dominates (2.36) for all values of $q'' \le \tilde{q}'$ and therefore $q'' = \tilde{q}'$ is the optimal choice. This can be seen from the following inequalities, for $(s-1)q'' + l > s\tilde{q}'$

$$3q'' - (r+1)\tilde{q}' - l + j \le -(r-1)\tilde{q}' \le -s\tilde{q}',$$

and for $(s-1)q'' + l \le s\tilde{q}'$

$$3q'' - (r+1)\tilde{q}' + j \le q'' - s\tilde{q}' + j \le q'' - (s-1)q'' - l + j \le -(s-1)q'',$$

where we assumed $r \ge s+1$. Using $l > j + \tilde{q}'$ and $\#\Lambda_{\mathbf{ja},l}^{(c)} = O(2^{3(l-j)})$, we obtain from (2.36) and (2.37) the final estimate

$$\sum_{\mathbf{b}\in\Lambda_{\mathbf{ja},l}^{(c)}} \left| \mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})} \right| \lesssim 2^{-s\tilde{q}'} 2^{-\frac{7}{2}j} 2^{-\frac{1}{2}l}. \tag{2.39}$$

In the second case with $l < j - \tilde{q}$ we take into account columns with $d_{\mathbf{ja},\star} \ge \eta 2^{-j+\tilde{q}}$ and the corresponding sets

$$\Lambda_{\mathbf{ja},l}^{(c')} := \left\{ \mathbf{b} \in \mathbb{B}_l : d_{\mathbf{ja},l\mathbf{b}} \le \eta 2^{-l} \text{ and } d_{l\mathbf{b},\star} \le \eta 2^{-l} \right\}. \tag{2.40}$$

Following the same line of arguments as before and using $\#\Lambda_{\mathbf{ja},l}^{(c')} = O(1)$, we obtain the estimate

$$\sum_{\mathbf{b}\in\Lambda_{\mathbf{ja},l}^{(c)}} \left| \mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})} \right| \lesssim 2^{-s\tilde{q}} 2^{-\frac{7}{2}j} 2^{-\frac{1}{2}l}. \tag{2.41}$$

Case D)
For columns with $d_{\mathbf{ja},\star} \le \eta 2^{-j}$ we consider the set

$$\Lambda_{\mathbf{ja},l}^{(d)} := \left\{ \mathbf{b} \in \mathbb{B}_l : d_{l\mathbf{b},\star} \le \eta 2^{-l+\tilde{q}'} \right\}.$$

Let us assume w.l.o.g. that both wavelets have a common nearest nucleus. With this we obtain for $l \geq j$ the estimate

$$\begin{aligned} \sum_{\mathbf{b}\in\Lambda^{(d)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| &\lesssim 2^{-j}2^{-\frac{7}{2}l} \sum_{\mathbf{b}\in\Lambda^{(d)}_{j\mathbf{a},l}} \left\|\partial_y\left(\phi(\mathbf{y})\int_{\mathbb{R}^3}\frac{\phi(\mathbf{x})\gamma_{j,\mathbf{a}}(\mathbf{x})}{|\mathbf{x}-\mathbf{y}|}d^3x\right)\right\|_{L_\infty(\square_{l,\mathbf{b}})} \\ &\lesssim 2^{3\tilde{q}'}2^{\frac{1}{2}j}2^{-\frac{7}{2}l}\int_0^{\eta 2^{-j}} du \\ &\lesssim 2^{3\tilde{q}'}2^{-\frac{1}{2}j}2^{-\frac{7}{2}l}, \end{aligned} \tag{2.42}$$

and the corresponding estimate for $l < j$ yields

$$\sum_{\mathbf{b}\in\Lambda^{(d)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{3\tilde{q}'}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.43}$$

Case E)
For $l \geq j$ we take into account columns with $d_{j\mathbf{a},\star} > \eta 2^{-j}$ and the corresponding sets

$$\Lambda^{(e)}_{j\mathbf{a},l} := \left\{\mathbf{b} \in \mathbb{B}_l : d_{j\mathbf{a},l\mathbf{b}} \leq \eta 2^{-j}\right\}, \tag{2.44}$$

where $d_{l\mathbf{b},\star} \geq \eta 2^{-j}$ for all $\mathbf{b} \in \Lambda^{(e)}_{j\mathbf{a},l}$ can be assumed.

Because of the absence of nuclei within the support of the wavelets, we can directly apply Proposition 2 and obtain the estimate

$$\begin{aligned} &\left|\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \gamma_{j,\mathbf{a}}(\mathbf{x})\frac{\phi(\mathbf{x})\phi(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\gamma_{l,\mathbf{b}}(\mathbf{y})d^3x d^3y\right| \\ &\qquad\lesssim 2^{-(s+\frac{5}{2})l}\left\|\partial_y^{s+1}\left(\phi(\mathbf{y})\int_{\mathbb{R}^3}\frac{\phi(\mathbf{x})\gamma_{j,\mathbf{a}}(\mathbf{x})}{|\mathbf{x}-\mathbf{y}|}d^3x\right)\right\|_{L_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} \\ &\qquad\lesssim 2^{-(s+\frac{5}{2})l}\max\left\{2^{(s-\frac{1}{2})j}, 2^{(s+\frac{1}{2})j}\right\} \\ &\qquad\lesssim 2^{-s(l-j)}2^{\frac{1}{2}j}2^{-\frac{5}{2}l}, \end{aligned}$$

where we have used for $0 < t \leq s+1$ the following estimate

$$\begin{aligned} \left\|\partial_y^t\int_{\mathbb{R}^3}\frac{\phi(\mathbf{x})\gamma_{j,\mathbf{a}}(\mathbf{x})}{|\mathbf{x}-\mathbf{y}|}d^3x\right\|_{L_\infty(\operatorname{supp}\gamma_{l,\mathbf{b}})} &\lesssim \left\|\partial_x^{t-1}(\phi\gamma_{j,\mathbf{a}})\right\|_{L_\infty(\operatorname{supp}\gamma_{j,\mathbf{a}})}\int_0^{\eta 2^{-j}} du \\ &\lesssim \sup_{t_1+t_2=t-1}\left\{\max\left\{2^{(t_1-1)j},1\right\}2^{(t_2+\frac{1}{2})j}\right\} \\ &\lesssim 2^{(t-\frac{1}{2})j}. \end{aligned}$$

Using $\#\Lambda^{(e)}_{j\mathbf{a},l} = O(2^{3(l-j)})$, we obtain for $l \geq j$

$$\sum_{\mathbf{b}\in\Lambda^{(e)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-s(l-j)}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.45}$$

Likewise for $l < j$, columns with $d_{j\mathbf{a},\star} > \eta 2^{-l}$ and sets

$$\Lambda^{(e')}_{j\mathbf{a},l} := \left\{\mathbf{b} \in \mathbb{B}_l : d_{j\mathbf{a},l\mathbf{b}} \leq \eta 2^{-l}\right\} \tag{2.46}$$

are taken into account, where $d_{l\mathbf{b},\star} \geq \eta 2^{-l}$ for all $\mathbf{b} \in \Lambda^{(e')}_{j\mathbf{a},l}$ can be assumed. From $\#\Lambda^{(e')}_{j\mathbf{a},l} = O(1)$, we obtain for $l < j$ the estimate

$$\sum_{\mathbf{b} \in \Lambda^{(e')}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-s(j-l)} 2^{-\frac{7}{2}j} 2^{-\frac{1}{2}l}. \tag{2.47}$$

Case F)
For columns with $d_{j\mathbf{a},\star} \geq \eta 2^{-j+\tilde{q}}$ we consider sets

$$\Lambda^{(f)}_{j\mathbf{a},l} := \left\{\mathbf{b} \in \mathbb{B}_l : d_{j\mathbf{a},l\mathbf{b}} \geq \eta 2^{-\min\{j,l\}+q'} \text{ and } d_{l\mathbf{b},\star} \leq \eta 2^{-l+\tilde{q}'}\right\}. \tag{2.48}$$

Using $\#\Lambda^{(f)}_{j\mathbf{a},l} = O(2^{3\tilde{q}'})$, we obtain for $l \geq j$ the estimate

$$\begin{aligned}
\sum_{\mathbf{b} \in \Lambda^{(f)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| &\lesssim 2^{-(r+\frac{5}{2})j} 2^{-\frac{7}{2}l} 2^{3\tilde{q}'} \left\| \partial^r_x \partial_y \left(\frac{\phi(\mathbf{x})\phi(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \right) \right\|_{L_\infty(\operatorname{supp}\gamma_{j,\mathbf{a}} \times \operatorname{supp}\gamma_{l,\mathbf{b}})} \\
&\lesssim 2^{-(r+\frac{5}{2})j} 2^{-\frac{7}{2}l} 2^{3\tilde{q}'} \sup_{t_1+t_2=r} \left\{ \max\left\{2^{(t_1-1)(j-\tilde{q})}, 1\right\} 2^{(t_2+2)(j-q')} \right\} \\
&\lesssim 2^{-\frac{1}{2}j} 2^{-\frac{7}{2}l} 2^{3\tilde{q}'} \max\left\{ 2^{-(r-1)\tilde{q}} 2^{-2q'} 2^{-j}, 2^{-(r+2)q'} \right\},
\end{aligned} \tag{2.49}$$

and for $l < j$ the corresponding estimate

$$\begin{aligned}
\sum_{\mathbf{b} \in \Lambda^{(f)}_{j\mathbf{a},l}} \left|\mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| &\lesssim 2^{-(r+\frac{5}{2})j} 2^{-\frac{7}{2}l} 2^{3\tilde{q}'} \left\| \partial^r_x \partial_y \left(\frac{\phi(\mathbf{x})\phi(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} \right) \right\|_{L_\infty(\operatorname{supp}\gamma_{j,\mathbf{a}} \times \operatorname{supp}\gamma_{l,\mathbf{b}})} \\
&\lesssim 2^{-(r+\frac{5}{2})j} 2^{-\frac{7}{2}l} 2^{3\tilde{q}'} \sup_{t_1+t_2=r} \left\{ \max\left\{2^{(t_1-1)(j-\tilde{q})}, 1\right\} 2^{(t_2+2)(l-q')} \right\} \\
&\lesssim 2^{-(r+\frac{5}{2})j} 2^{-\frac{7}{2}l} 2^{3\tilde{q}'} \max\left\{ 2^{(r-1)(j-\tilde{q})} 2^{2(l-q')}, 2^{(r+2)(l-q')} \right\}.
\end{aligned} \tag{2.50}$$

2.6 Proof of Lemma 3

In the first part of the proof, we assume that the adaptive refinement scheme presented in Lemma 3 has been applied to arbitrarily fine levels, i.e. $j_{max} = \infty$. This is in conflict with our requirement that the compressed matrix has at most $O(2^{3q})$ nonzero entries in each row and column. The following result therefore requires further improvement.

Proposition 4. *The spectral norm of the residuum matrix, for $j_{max} = \infty$, with entries*

$$\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})} := \mathcal{U}_{(j,\mathbf{a}),(l,\mathbf{b})} - \mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})} \tag{2.51}$$

is bounded by the estimate

$$\left\| \{\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\}_{(j,\mathbf{a}),(l,\mathbf{b}) \in \Lambda} \right\| \lesssim 2^{-(s-\epsilon)q}, \tag{2.52}$$

for any $q \in \mathbb{N}$ and $\epsilon > 0$, where the wavelet basis satisfies (2.2) and (2.4) with $r \geq s+1$ vanishing moments.

Proof. In order to estimate the error of the residuum matrix (2.51) in the spectral norm, we apply Schur's lemma for symmetric matrices, cf. Lemma 1, using the variant discussed in Ref. [32]. This enables a subdivision into estimates for different blocks of the residuum matrix which can be derived successively. According to the adaptive refinement scheme of Lemma 3 we first consider the case $j+q<l$. In this case nonzero entries in a column of the residuum matrix are restricted to the index sets (2.24), (2.30), (2.35), (2.44) and (2.48). The union of the first four index sets, with $\tilde{q}'=q$ and $q'=0$, is given by

$$\Lambda^{(a)}_{\mathbf{ja},l}\cup\Lambda^{(b)}_{\mathbf{ja},l}\cup\Lambda^{(c)}_{\mathbf{ja},l}\cup\Lambda^{(e)}_{\mathbf{ja},l}=\left\{\mathbf{b}\in\mathbb{B}_l : d_{l\mathbf{b},\star}\geq\eta 2^{-l+q}\right\},$$

which however require separate estimates.

Case A and B where (2.26) and (2.32) apply with $\tilde{q}'=q$ and $q'=0$

$$\sum_{\mathbf{b}\in\Lambda^{(a)}_{\mathbf{ja},l}\cup\Lambda^{(b)}_{\mathbf{ja},l}}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right|\lesssim 2^{-(r-1)q}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.53}$$

Case C where (2.39) applies with $\tilde{q}'=q$ and $q'=0$

$$\sum_{\mathbf{b}\in\Lambda^{(c)}_{\mathbf{ja},l}}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right|\lesssim 2^{-sq}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.54}$$

Case E where (2.45) applies with $q'=0$

$$\sum_{\mathbf{b}\in\Lambda^{(e)}_{\mathbf{ja},l}}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right|\lesssim 2^{-s(l-j)}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}\leq 2^{-sq}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.55}$$

Case F refers to the index set (2.48) with $q'=\tilde{q}=\tilde{q}'=q$, and estimate (2.49) yields

$$\begin{aligned}\sum_{\mathbf{b}\in\Lambda^{(f)}_{\mathbf{ja},l}}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| &\lesssim 2^{-\frac{1}{2}j}2^{-\frac{7}{2}l}2^{3q}\max\left\{2^{-(r+1)q}2^{-j},2^{-(r+2)q}\right\}\\ &\lesssim 2^{-(r-1)q}2^{-\frac{1}{2}j}2^{-\frac{7}{2}l}.\end{aligned} \tag{2.56}$$

From Schur's lemma together with the estimates (2.53), (2.55), (2.54) and (2.56) for $r=s+1$, we get the block estimate

$$\begin{aligned}\sup_{j\geq 0}\sum_{l>j+q}\sum_{\mathbf{b}\in\mathbb{B}_l}2^{p(j-l)}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| &\lesssim \sup_{j\geq 0}\sum_{l>j+q}2^{-sq}2^{(p-\frac{7}{2})j}2^{-(p+\frac{1}{2})l}\\ &\lesssim 2^{-(s+p+\frac{1}{2})q}\quad\text{for } p>-\tfrac{1}{2}.\end{aligned} \tag{2.57}$$

On the next level of adaptivity additional diagonal refinement has been taken into account. For $j\leq l\leq j+q$ nonzero entries in a column of the residuum matrix are restricted to the index sets (2.24), (2.30) and (2.48).

Case A and B apply with $\tilde{q}=0$, $\tilde{q}'=q$ and $q'=q-l+j$. The union of the index sets (2.24), (2.30) is given by

$$\Lambda^{(a)}_{\mathbf{ja},l}\cup\Lambda^{(b)}_{\mathbf{ja},l}=\left\{\mathbf{b}\in\mathbb{B}_l : d_{\mathbf{ja},l\mathbf{b}}\geq\eta 2^{-l+q}\text{ and } d_{l\mathbf{b},\star}\geq\eta 2^{-l+q}\right\},$$

for which (2.27) and (2.33) yield the estimate

$$\sum_{\mathbf{b}\in\Lambda^{(a)}_{\mathbf{ja},l}\cup\Lambda^{(b)}_{\mathbf{ja},l}}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right|\lesssim 2^{-(r-1)q}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.58}$$

Case F applies with $q' = \tilde{q} = \tilde{q}' = q$ and index set (2.48) for which (2.49) yields the estimate

$$\sum_{\mathbf{b}\in\Lambda^{(f)}_{j\mathbf{a},l}} \left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-(r-1)q}2^{-\frac{1}{2}j}2^{-\frac{7}{2}l}. \tag{2.59}$$

In both cases, the estimate is valid for $r \geq s+1$. Using Schur's lemma and (2.58), (2.59) we obtain the block estimate

$$\begin{aligned}\sup_{j\geq 0}\sum_{j\leq l\leq j+q}\sum_{\mathbf{b}\in\mathbb{B}_l} 2^{p(j-l)}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| &\lesssim \sup_{j\geq 0}\sum_{j\leq l\leq j+q} 2^{-(r-1)q}2^{(p-\frac{7}{2})j}2^{-(p+\frac{1}{2})l} \\ &\lesssim 2^{-(r-1)q} \quad \text{for } p > -\tfrac{1}{2}\end{aligned} \tag{2.60}$$

Correspondingly, for $j-q \leq l < j$ nonzero entries in a column of the residuum matrix are restricted to the index sets (2.24), (2.30) and (2.48).

Case A and F apply with $\tilde{q} = q$, $\tilde{q}' = 0$ and $q' = q-j+l$. The union of the index sets (2.24) and (2.48), with $d_{j\mathbf{a},\star} \geq \eta 2^{-j+q}$, is given by

$$\Lambda^{(a)}_{j\mathbf{a},l} \cup \Lambda^{(f)}_{j\mathbf{a},l} = \left\{\mathbf{b}\in\mathbb{B}_l : d_{j\mathbf{a},l\mathbf{b}} \geq \eta 2^{-j+q}\right\},$$

for which (2.28) and (2.50) yield the estimate

$$\sum_{\mathbf{b}\in\Lambda^{(a)}_{j\mathbf{a},l}\cup\Lambda^{(f)}_{j\mathbf{a},l}} \left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-(r-1)q}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.61}$$

Case B applies with $q' = \tilde{q} = \tilde{q}' = q$ and index set (2.30) for which (2.34) yields the estimate

$$\sum_{\mathbf{b}\in\Lambda^{(b)}_{j\mathbf{a},l}} \left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-(r-1)q}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l} \tag{2.62}$$

Again in both cases the estimate is valid for $r \geq s+1$. Using Schur's lemma and (2.61), (2.62) we obtain the block estimate

$$\begin{aligned}\sup_{j\geq 0}\sum_{j-q\leq l<j}\sum_{\mathbf{b}\in\mathbb{B}_l} 2^{p(j-l)}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| &\lesssim \sup_{j\geq 0}\sum_{j-q\leq l\leq j} 2^{-(r-1)q}2^{(p-\frac{7}{2})j}2^{-(p+\frac{1}{2})l} \\ &\lesssim 2^{-(r-p-\frac{3}{2})q} \quad \text{for } p > -\tfrac{1}{2}\end{aligned} \tag{2.63}$$

In the last case $0 \leq l < j-q$, nonzero entries in a column of the residuum matrix are restricted to the index sets (2.24), (2.40), (2.46), (2.48) and (2.30). The union of the first four index sets, with $\tilde{q} = q$ and $q' = \tilde{q}' = 0$, is given by

$$\Lambda^{(a)}_{j\mathbf{a},l} \cup \Lambda^{(c')}_{j\mathbf{a},l} \cup \Lambda^{(e')}_{j\mathbf{a},l} \cup \Lambda^{(f)}_{j\mathbf{a},l} = \left\{\mathbf{b}\in\mathbb{B}_l : d_{j\mathbf{a},\star} \geq \eta 2^{-j+q}\right\},$$

which however requires two separate estimates.

Case A and F where (2.29) and (2.50) apply with $\tilde{q} = q$ and $q' = \tilde{q}' = 0$

$$\sum_{\mathbf{b}\in\Lambda^{(a)}_{j\mathbf{a},l}\cup\Lambda^{(f)}_{j\mathbf{a},l}} \left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-(r-1)q}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.64}$$

Case B refers to the index set (2.30) with $q' = \tilde{q} = \tilde{q}' = q$ and estimate (2.34) yields

$$\sum_{\mathbf{b}\in\Lambda^{(b)}_{j\mathbf{a},l}} \left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-(r-1)q}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l} \tag{2.65}$$

Case C and E where (2.41) and (2.47) apply with $\tilde{q} = q$ and $q' = \tilde{q}' = 0$

$$\sum_{\mathbf{b}\in\Lambda^{(c')}_{j\mathbf{a},l}\cup\Lambda^{(e')}_{j\mathbf{a},l}} \left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-sq}2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.66}$$

From Schur's lemma and estimates (2.64), (2.65), (2.66) with $r = s+1$ we get the block estimate

$$\begin{aligned}\sup_{j\geq 0}\sum_{0\leq l<j-q}\sum_{\mathbf{b}\in\mathbb{B}_l} 2^{p(j-l)}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| &\lesssim \sup_{j\geq 0}\sum_{0\leq l<j-q} 2^{-sq}2^{(p-\frac{7}{2})j}2^{-(p+\frac{1}{2})l}\\ &\lesssim 2^{-sq} \quad \text{for } -\tfrac{1}{2} < p \leq \tfrac{7}{2}.\end{aligned} \tag{2.67}$$

Eventually, we have to combine the block estimates to get an estimate for the spectral norm of the residuum matrix. Let us consider the case $r = s+1$ where we use the block estimates (2.57), (2.60), (2.63) and (2.67). A perfectly balanced estimate for the p dependent cases (2.57) and (2.63) demands

$$2^{-(s+p+\frac{1}{2})q} \sim 2^{-(s-p-\frac{1}{2})q}$$

which is satisfied for $p = -\frac{1}{2}$. Since the block estimates require $-\frac{1}{2} < p \leq \frac{7}{2}$ at best $p = -\frac{1}{2} + \epsilon$ for any $0 < \epsilon \leq 4$ can be achieved. We can now apply Schur's lemma to estimate the error of the residuum matrix (2.51) in the spectral norm

$$\left\|\{\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\}_{(j,\mathbf{a}),(l,\mathbf{b})\in\Lambda}\right\| \lesssim \sup_{(j,\mathbf{a})\in\Lambda}\sum_{(l,\mathbf{b})\in\Lambda} 2^{-(\frac{1}{2}-\epsilon)(j-l)}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-(s-\epsilon)q},$$

from which (2.52) follows. □

It is obvious, cf. Fig. 1, that we cannot maintain adaptive refinements for arbitrarily fine wavelet levels j, l while preserving our bounds for the number of nonzero entries per row and column. Therefore a q dependent maximum level j_{max} is required which provides a bound on the number of rows and columns containing nonzero entries.

Corollary 1. *The matrix* $\{\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\}_{(j,\mathbf{a}),(l,\mathbf{b})\in\Lambda}$ *can be further compressed by neglecting all matrix entries with index* $j \geq j_{max}$ *or* $l \geq j_{max}$*, at which* $j_{max} \simeq \frac{s-\epsilon}{4-\epsilon}q$*, with* $0 < \epsilon < 4$*, has been chosen so that the estimate (2.52) remains valid.*

Proof. Collecting all previous estimates for the cases A to F, with $q' = \tilde{q} = \tilde{q}' = 0$, together we obtain

$$\sum_{\mathbf{b}\in\mathbb{B}_l} \left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-\frac{7}{2}j}2^{-\frac{1}{2}l}. \tag{2.68}$$

For $l \geq j$ this yields

$$\sup_{j\geq j_{max}}\sum_{l\geq j}\sum_{\mathbf{b}\in\mathbb{B}_l} 2^{p(j-l)}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{-4j_{max}}, \tag{2.69}$$

and for $l < j$

$$\sup_{j\geq j_{max}}\sum_{l<j}\sum_{\mathbf{b}\in\mathbb{B}_l} 2^{p(j-l)}\left|\Delta\mathcal{U}^{(q)}_{(j,\mathbf{a}),(l,\mathbf{b})}\right| \lesssim 2^{(p-\frac{7}{2})j_{max}}, \tag{2.70}$$

with $-\frac{1}{2} < p < \frac{7}{2}$. Like in the previous proposition, we take $p = -\frac{1}{2} + \epsilon$ for $0 < \epsilon < 4$ and demand

$$2^{-(4-\epsilon)j_{max}} \sim 2^{-(s-\epsilon)q},$$

in order to obtain a balanced estimate between (2.52) and (2.70). This can be achieved by requiring

$$j_{max} \simeq \frac{s-\epsilon}{4-\epsilon} q.$$

□

Proof of Lemma 3

Proof. The number of nonzero entries per row and column is bounded by a constant multiple of $q2^{3q} \lesssim 2^{(3+\epsilon)q}$ which must be compared with the spectral error (2.52) of the residuum matrix. For any $0 < \alpha < s^*$, we can therefore choose an $\epsilon > 0$ such that

$$\alpha \leq \frac{s-\epsilon}{3+\epsilon},$$

and the lemma follows from immediately from Definition 1. □

3 Acknowledgments

The authors gratefully acknowledge Andreas Zeiser (Berlin) for useful discussions. This work was supported by the Deutsche Forschungsgemeinschaft (Grant No. SCHN 530/5-2).

References

[1] Andrae, D.: Numerical self-consistent field method for polyatomic molecules. Mol. Phys. **99**, 327-334 (2001).

[2] Arias, T. A.: Multiresolution analysis of electronic structure: Semicardinal and wavelet bases. Rev. Mod. Phys. **71**, 267–312 (1999).

[3] Beck, O., Heinemann, D., Kolb, D.: Fast and accurate molecular Hartree-Fock with a finite-element multigrid method. arXiv:physics/0307108, pp. 20 (2003).

[4] Braess, D.: Asymptotics for the approximation of wave functions by exponential sums. J. Approx. Theory **83**, 93-103 (1995).

[5] Bungartz, H.-J., Griebel, M.: Sparse grids, Acta Numerica **13**, 147-269 (2004).

[6] Cancès E.: SCF algorithms for HF electronic calculations. In Defranceschi M., Le Bris C. (eds.). *Mathematical Models and Methods for Ab Initio Quantum Chemistry, Lecture Notes in Chemistry* **74**. Springer: Berlin: 2000, 17-43.

[7] Cancès E., Le Bris C.: On the convergence of SCF algorithms for the Hartree-Fock equations. ESAIM: M2AN **34**, 749–774 (2000).

[8] Cohen, A., Dahmen, W., DeVore, R. A.: Adaptive wavelet methods for elliptic operator equations, convergence rates. Math. Comp. **70**, 27-75 (2001).

[9] Dahmen, W., Rohwedder, T., Schneider R., Zeiser, A.: Adaptive eigenvalue computation - complexity estimates: to appear in Numer. Math.

[10] DeVore, R. A.: Nonlinear approximation. Acta Numerica **7**, 51-150 (1998).

[11] Egorov Y. V., Schulze B.-W.: Pseudo-Differential Operators, Singularities, Applications. Basel: Birkhäuser 1997.

[12] Engeness, T. D., Arias, T. A.: Multiresolution analysis for efficient, high precision all-electron density-functional calculations. Phys. Rev. B **65**, 165106, pp. 10 (2002).

[13] Flad, H.-J., Hackbusch, W., Luo H., Kolb, D.: Diagrammatic multiresolution analysis for electron correlations. Phys. Rev. B. **71**, 125115, 18 p. (2005).

[14] Flad, H.-J., Hackbusch, W., Schneider, R.: Best N-term approximation in electronic structure calculation. I. One-electron reduced density matrix. ESAIM: M2AN **40**, 49–61 (2006).

[15] Flad, H.-J., Schneider, R., Schulze, B.-W.: Asymptotic regularity of solutions of Hartree-Fock equations with Coulomb potential. to appear in Math. Methods Appl. Sci.

[16] Genovese, L., Deutsch, T., Neelov, A., Goedecker, S., Beylkin, G.: Efficient solution of Poisson's equation with free boundary conditions. J. Chem. Phys. 125, 074105, pp. 5 (2006).

[17] Genovese, L., Neelov, A., Goedecker, S., Deutsch, T., Ghasemi, S. A., Willand, A., Caliste, D., Zilberberg, O., Rayson, M., Bergman, A., Schneider, R.: Daubechies wavelets as a basis set for density functional pseudopotential calculations. J. Chem. Phys. **129**, 014109, pp. 14 (2008).

[18] Gilbarg, D., Trudinger, N. S.: Elliptic Partial Differential Equations of Second Order. Berlin: Springer 1998.

[19] Heinemann, D. Rosén, A., Fricke B.: Solution of the Hartree-Fock equations for atoms and diatomic molecules with the finite element method. Phys. Scripta **42**, 692-696 (1990).

[20] Harrison, R. J., Fann, G.I., Yanai, T., Gan, Z., Beylkin, G.: Multiresolution quantum chemistry: Basic theory and initial applications. J. Chem. Phys. **121**, 11587-11598 (2004).

[21] Helgaker, T., Jørgensen, P., Olsen, J.: Molecular Electronic-Structure Theory. New York: Wiley 1999.

[22] Kobus, J., Laaksonen, L., Sundholm, D.: A numerical Hartree-Fock program for diatomic molecules. Comput. Phys. Commun. **98**, 346-358 (1996).

[23] Kutzelnigg, W.: Theory of the expansion of wave functions in a Gaussian basis. Int. J. Quantum Chem. **51**, 447-463 (1994).

[24] Lieb, E. H., Simon, B.: The Hartree-Fock theory for Coulomb systems. Commun. Math. Phys. **53**, 185-194 (1977).

[25] Lions, P. L.: Solutions of Hartree-Fock equations for Coulomb systems. Commun. Math. Phys. **109**, 33-97 (1987).

[26] Neelov, A. I., Goedecker, S.: An efficient numerical quadrature for the calculation of the potential energy of wavefunctions expressed in the Daubechies wavelet basis. J. Comp. Phys. **217**, 312-339 (2006).

[27] Rohwedder, T., Schneider, R., Zeiser, A.: Perturbed preconditioned inverse iteration for operator eigenvalue problems with application to adaptive wavelet discretization. submitted to Adv. in Comp. Math.

[28] Sinanoğlu, O.: Perturbation theory of many-electron atoms and molecules. Phys. Rev. **122**, 493-499 (1961).

[29] Sinanoğlu, O.: Theory of electron correlation inatoms and molecules. Proc. R. Soc. London, Ser. A **260**, 379-392 (1961).

[30] Stevenson, R.: On the compressibility of operators in wavelet coordinates. SIAM J. Math. Anal. **35**, 1110–1132 (2004).

[31] Schwab, C., Stevenson, R.: Adaptive wavelet algorithms for elliptic PDE's on product domains. Research Report 2006-16 ETH Zürich.

[32] Schneider R.: Multiskalen- und Wavelet-Matrixkompression. Stuttgart: Teubner 1998.

[33] Yanai, T., Fann, G. I., Gan, Z., Harrison, R. J., Beylkin, G.: Multiresolution quantum chemistry in multiwavelet basis: Hartree-Fock exchange. J. Chem. Phys. **121**, 6680-6688 (2004).

[34] Yserentant, H., On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives, Numer. Math. **98**, 731-759 (2004).