Aachen Institute for Advanced Study in Computational Engineering Science

Preprint: AICES-2010/06-2

18/June/2010

Hyperbolic Wavelet Discretization of the Two-Electron Schrödinger Equation in an Explicitly Correlated Formulation

M. Bachmayr

RWTH AACHEN UNIVERSITY

AICES

Financial support from the Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111 is gratefully acknowledged.

List of AICES technical reports: http://www.aices.rwth-aachen.de/preprints

Hyperbolic Wavelet Discretization of the Two-Electron Schrödinger Equation in an Explicitly Correlated Formulation

Markus Bachmayr*

Abstract

In the framework of an explicitly correlated formulation of the electronic Schrödinger equation known as the transcorrelated method, this work addresses some fundamental issues concerning the feasibility of eigenfunction approximation by hyperbolic wavelet bases. Focusing on the two-electron case, the integrability of mixed weak derivatives of eigenfunctions of the modified problem is investigated, obtaining a significant improvement compared to the standard formulation. Elements of a discretization of the eigenvalue problem based on orthogonal wavelets, in which separable approximations of potential terms are used to apply operators efficiently, are described, and estimates for the error due to this further approximation are given.

Mathematics Subject Classification (2010): 35B65, 35J10, 65T60, 81Q05.

1 Introduction

The Schrödinger equation is the basic equation of non-relativistic quantum physics. In its time-dependent form, it describes the time evolution of quantum states; in the stationary case, which we shall be dealing with, it becomes an eigenvalue problem which has as its solutions the possible quantum states that the physical system can attain.

For the description of molecular systems, it is usually sufficient to only describe the electrons by the Schrödinger equation and model the nuclei as classical particles. Assuming a system of n electrons and some given nuclei, indexed by ν, of charges Z_ν clamped at positions $a_\nu \in \mathbb{R}^3$, this so-called Born-Oppenheimer approximation leads to the electronic Schrödinger equation for the wavefunction $u\colon (\mathbb{R}^3 \times \{-1/2, 1/2\})^n \to \mathbb{R}$,

$$\left\{ -\frac{1}{2} \sum_{i=1}^{n} \Delta_{x_i} - \sum_{i=1}^{n} \sum_{\nu} \frac{Z_\nu}{|x_i - a_\nu|} + \sum_{1 \le i < j \le n} \frac{1}{|x_i - x_j|} \right\} u = \lambda u \,. \tag{1.1}$$

For each electron, we have coordinates $(x, \sigma) \in \mathbb{R}^3 \times \{-1/2, 1/2\}$, where the spin variable σ enters only via an additional antisymmetry requirement on u: the wavefunction needs to change sign when exchanging (x_i, σ_i) and (x_j, σ_j) for any $i \neq j$. Due to the symmetry properties of the Hamiltonian, the problem reduces to solving (1.1) on $\mathbb{R}^{3n}$ for $\lfloor n/2 \rfloor$ different fixed spin configurations, where each of the resulting spatial components needs to be antisymmetric under exchange of spatial coordinates of electrons that have equal spin; in what follows, by a solution of (1.1) we always mean such a spatial component for a certain spin configuration.

Without the two-electron interaction potentials $|x_i - x_j|^{-1}$, these solutions would be antisymmetrized products of single-electron orbitals, i.e. of functions on $\mathbb{R}^3$. This is the approximation made in the Hartree-Fock method, which gives surprisingly good results for many systems, but only gives strict upper bounds for the energies λ in the presence of two-electron interaction terms.

*RWTH Aachen, Aachen Institute for Advanced Study in Computational Engineering Sciences, Schinkelstr. 2, 52062 Aachen, Germany (bachmayr@aices.rwth-aachen.de)

Post-Hartree-Fock methods such as configuration interaction or coupled cluster aim to improve the Hartree-Fock approximation by combining several antisymmetrized products. Although these standard methods based on expansions into products of one-electron orbitals can give good first approximations, the actual convergence towards the exact solution is very slow. The main reason is that the solution is nonsmooth at singularities of potentials. The cusps arising from the nuclear potentials are quantitatively stronger, but can be approximated very efficiently using standard methods based on Gaussian-type basis functions, whereas the correlation cusps at two-electron coalescence points turn out to be the most problematic aspect.

The simplest system where this can be observed is Helium, consisting of one nucleus of charge $Z = 2$ and two electrons, where (1.1) becomes

$$-\frac{1}{2}\Delta u - \frac{2}{|x_1|}u - \frac{2}{|x_2|}u + \frac{1}{|x_1 - x_2|}u = \lambda u\,. \tag{1.2}$$

The first-order behavior at the cusp was characterized by [19]; in the form given in [15] it amounts to the assertion that eigenfunctions u of (1.2) can be written as

$$u(x) = \exp\Big(-2|x_1| - 2|x_2|\Big)\exp\Big(\frac{1}{2}|x_1 - x_2|\Big)w(x)$$

with $w \in C^{1,\alpha}(\mathbb{R}^{3n})$ for $\alpha \in (0,1)$ – in other words, in the vicinity of two-electron coalescence points, u looks essentially like $1 + \frac{1}{2}|x_1 - x_2|$. It was shown in [14] that approximation of such a function by a sum of N products of one-electron functions can at best give an H^1-error that decays almost as $N^{-1/2}$, and available Post-Hartree-Fock methods based on this type of approximation actually perform worse.

Methods that aim to improve on this situation have a long history, starting with the work of Hylleraas in 1929 [18], who provided the first accurate calculations of the ground state of Helium. Exploiting symmetries of this particular situation, he used an expansion of the form

$$\psi \approx \exp\big(-\zeta s\big)\sum_{i,j,k} C_{ijk}\, s^i\, t^{2j}\, u^k$$

with coordinates $s = |x_1| + |x_2|$, $t = |x_1| - |x_2|$, $u = |x_1 - x_2|$ and unknowns ζ and C_{ijk}. Despite the excellent results this method yields for Helium and, in a much more complicated form, for three- or four-electron atomic systems, the approach is very difficult to adapt to more electrons or more complicated molecular geometries.

It was observed, however, that including terms that are linear in $|x_1 - x_2|$ into standard orbital expansions already yields significantly improved results, which leads to the class of R12 methods. A variety of such explicitly correlated schemes has been developed, see e.g. [20] for an overview. The factors depending on $|x_1 - x_2|$ that are included in the expansion are referred to as correlation factors or Jastrow factors in this context; when combined with standard methods that use a small set of Gaussian-type basis functions, the particular choice of such factors can have considerable influence on the quality of the results, cf. [28].

Despite the success and versatility of quantum chemical methods based on Gaussian basis sets, their mathematical understanding remains limited, and they do not offer a systematic way of controlling errors and refining approximations. Recent eigenfunction regularity results of Yserentant [29, 31] show that it is in principle possible to obtain a discretization with controlled error based on sparse grid- or hyperbolic wavelet-type approximation schemes. For practically relevant problems, there is a lot of work to be done before such generic methods can be expected to be competitive with optimized, problem-specific Gaussian methods, but they offer a starting point for electronic structure methods that facilitate a rigorous analysis. The objective of this study is to advance in this direction.

The results collected in [31] show that n-electron eigenfunctions of (1.1) can be approximated at essentially the cost corresponding to a full tensor product discretization of a two-electron problem, where the obstacle for a further reduction of the complexity is the restriction of regularity

due to the electron-electron cusp. In the present work, we use an explicitly correlated ansatz to arrive at a modified problem where the electron-electron singularities are eliminated. We adapt the basic argument of [29] to quantify the resulting improvement in regularity of solutions, which allows to reduce the complexity of representing a two-electron wavefunction using sparse tensor products of wavelets down to almost that of a one-electron problem, hence approaching practical feasibility.

In Section 2, we discuss a particular explicitly correlated formulation known in the computational chemistry literature as the transcorrelated method. In Section 3, a regularity result for the correspondingly modified eigenfunctions in the two-electron case is proven, and some basic consequences for approximation by wavelets are discussed in Section 4. The actual aim is of course not only to represent a solution, but to solve an eigenvalue problem; therefore in Section 5, we discuss some important building blocks for actually computing such an approximation.

After completion of this work, I became aware of a new compilation of results by Yserentant [30] that may employ similar concepts.

2 Transcorrelated Formulation

In this section, we look at a specific way of incorporating information on the electron-electron cusp into approximations of wavefunctions: The electron-electron singularities are eliminated at the price of introducing additional nonsymmetric first-order two-electron terms and symmetric zero-order three-electron terms, whereas the single-electron parts of the Hamiltonian are unchanged.

For the weak formulation of the eigenvalue problem, we define the bilinear form a for $u, v \in H^1(\mathbb{R}^{3n})$ as

$$a(u,v) := \frac{1}{2}\int \nabla u \cdot \nabla v \, dx + \int (V_{\mathrm{ne}} + V_{\mathrm{ee}}) uv \, dx \tag{2.1}$$

with the potential terms

$$V_{\mathrm{ne}}(x) := -\sum_{i,\nu} \frac{Z_\nu}{|x_i - a_\nu|}\,, \quad V_{\mathrm{ee}}(x) := \sum_{i<j} \frac{1}{|x_i - x_j|}\,.$$

We say that $u \in H^1(\mathbb{R}^{3n})$, $u \neq 0$ is an eigenfunction of a with eigenvalue λ if

$$a(u,v) = \lambda \langle u, v\rangle \quad \text{for all } v \in H^1(\mathbb{R}^{3n}), \tag{2.2}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing induced by the L^2-inner product.

In explicitly correlated methods, an ansatz for u is made that explicitly includes the correct first-order behavior of the electron-electron cusp, and v is chosen accordingly to obtain a favorable modified bilinear form. For the further discussion, let

$$F(x) := \frac{1}{2}\sum_{i<j} |x_i - x_j|\,.$$

A first option, related to the mentioned R12 methods, would be an ansatz $u = (1+F)\varphi$, $v = (1+F)\tau$. This preserves symmetry of the Hamiltonian, but leads to rather complicated four-electron integrals.

The approach we will follow here corresponds to taking

$$u = \exp(F)\varphi\,, \quad v = \exp(-F)\tau$$

instead, which can be interpreted as a similarity transformation and in the computational chemistry literature is referred to as a transcorrelated method. Although it entails loss of symmetry of

the bilinear form, it completely eliminates the two-electron singularities and avoids four-electron integrals[1].

For $\varphi, \tau \in H^1(\mathbb{R}^{3n})$, the modified bilinear form is given by

$$\tilde{a}(\varphi,\tau) := \frac{1}{2}\int \nabla\varphi\cdot\nabla\tau\,dx + \int V_{\mathrm{ne}}\varphi\tau\,dx - \int(\nabla F\cdot\nabla\varphi)\tau\,dx - \frac{1}{2}\int|\nabla F|^2\varphi\tau\,dx\,.$$

Written out in full, we have

$$\begin{aligned}\tilde{a}(\varphi,\tau) = \frac{1}{2}\int\nabla\varphi\cdot\nabla\tau\,dx - \sum_{i,\nu}\int\frac{Z_\nu}{|x_i-a_\nu|}\varphi\,\tau\,dx \\ - \frac{1}{2}\sum_i\int\sum_{k\neq i}\frac{x_i-x_k}{|x_i-x_k|}\cdot\nabla_{x_i}\varphi\,\tau\,dx - \frac{1}{8}\sum_i\int\sum_{\substack{k\neq i\\ l\neq i}}\frac{x_i-x_k}{|x_i-x_k|}\cdot\frac{x_i-x_l}{|x_i-x_l|}\,\varphi\,\tau\,dx\,. \quad (2.3)\end{aligned}$$

The strong form of the modified problem (2.3) was also used to obtain the regularity results in [15] already mentioned above. In the quantum chemistry literature, the formulation seems to appear first in [17]; it was used in computational schemes for Gaussian-type orbitals for instance in [3, 23, 27, 32]. Quite promising numerical results using Gaussian basis sets for Helium for the particular Hamiltonian corresponding to (2.3) are given in [20].

We begin by establishing the connection between a and $\tilde{a}$.

Proposition 2.1. *Let $u \in H^1(\mathbb{R}^{3n})$ be an eigenfunction of a with eigenvalue λ, then $w = e^{-F}u \in H^1(\mathbb{R}^{3n})$ is an eigenfunction of $\tilde{a}$ with the same eigenvalue,*

$$\tilde{a}(w,\tau) = \lambda\langle w,\tau\rangle \quad \textit{for all } \tau \in H^1(\mathbb{R}^{3n}). \quad (2.4)$$

Proof. From Lipschitz continuity of $\exp(-|\cdot|)$ and the chain and product rules for weak differentiation (cf. [16]) it follows that $e^{-F}u \in H^1(\mathbb{R}^{3n})$. For any $\tau \in \mathcal{D}(\mathbb{R}^{3n})$,

$$\begin{aligned}\tilde{a}(e^{-F}u,\tau) &= \frac{1}{2}\int e^{-F}(\nabla u - u\nabla F)\cdot\nabla\tau\,dx - \int e^{-F}\nabla F\cdot(\nabla u - u\nabla F)\tau\,dx \\ &\qquad - \frac{1}{2}\int e^{-F}|\nabla F|^2 u\tau\,dx + \int e^{-F}V_{\mathrm{ne}}u\tau\,dx \\ &= \frac{1}{2}\int e^{-F}\nabla u\cdot\nabla\tau\,dx - \frac{1}{2}\int e^{-F}u\nabla F\cdot\nabla\tau\,dx - \int e^{-F}\nabla F\cdot\nabla u\,\tau\,dx \\ &\qquad + \frac{1}{2}\int e^{-F}|\nabla F|^2 u\tau\,dx + \int e^{-F}V_{\mathrm{ne}}u\tau\,dx\,.\end{aligned}$$

Now on the one hand,

$$\frac{1}{2}\int e^{-F}\nabla u\cdot\nabla\tau\,dx - \frac{1}{2}\int e^{-F}\nabla F\cdot\nabla u\,\tau\,dx = \frac{1}{2}\int\nabla u\cdot\nabla(e^{-F}\tau)\,dx\,,$$

on the other hand by integration by parts and noting that $\Delta F = 2V_{\mathrm{ee}}$,

$$\begin{aligned}-\frac{1}{2}\int e^{-F}u\nabla F\cdot\nabla\tau\,dx &= \frac{1}{2}\int\nabla(e^{-F}u\nabla F)\tau\,dx \\ &= \frac{1}{2}\int\nabla(e^{-F}u)\cdot\nabla F\,\tau\,dx + \frac{1}{2}\int e^{-F}\Delta F\,u\tau\,dx \\ &= \frac{1}{2}\int e^{-F}(\nabla u\cdot\nabla F)\tau\,dx - \frac{1}{2}\int e^{-F}|\nabla F|^2 u\tau\,dx + \int e^{-F}V_{\mathrm{ee}}u\tau\,dx\,.\end{aligned}$$

[1]It is also possible to modify the correlation factor $\exp(F)$ to a uniformly bounded function and still achieve the same effect. This is important for modelling the correct decay behaviour when using Gaussian basis sets, but since this is not an issue in our setting we use the unbounded correlation factor $\exp(F)$, which can be considered a limiting case and leads to less complicated expressions.

Putting this together, using that $e^{-F}\tau \in H^1(\mathbb{R}^{3n})$ and that u solves (2.2),

$$\begin{aligned}\tilde{a}(e^{-F}u,\tau) &= \frac{1}{2}\int \nabla u\cdot\nabla(e^{-F}\tau)\,dx + \int e^{-F}(V_{\mathrm{ne}}+V_{\mathrm{ee}})u\tau\,dx\\ &= a(u,e^{-F}\tau) = \lambda\langle u,e^{-F}\tau\rangle = \lambda\langle e^{-F}u,\tau\rangle\,,\end{aligned}$$

and by density we obtain the assertion. □

As mentioned, the disadvantage of the modified problem for $\tilde{a}$ is that the symmetry of the bilinear form a is lost. This has the further consequence that for an eigenfunction u with eigenvalue λ of a, the solutions of the adjoint problem

$$\tilde{a}(\tau,w^*) = \lambda\langle w^*,\tau\rangle \quad \text{for all } \tau\in H^1(\mathbb{R}^{3n}) \tag{2.5}$$

are different from those of (2.4): It will be shown below that (2.5) is solved by $w^* = e^F u$, given that this product is contained in $H^1(\mathbb{R}^{3n})$. Whereas $e^F u \in H^1_{\mathrm{loc}}(\mathbb{R}^{3n})$ follows as in the proof of Proposition 2.1 from $u\in H^1(\mathbb{R}^{3n})$, global square integrability now depends on the decay properties of u.

The decay behavior of eigenfunctions of Schrödinger operators has been studied quite extensively, see the monograph [1]. We quote a result on exponential decay in the L^2-sense given in [31] that is most appropriate for our purposes.

Theorem 2.2. *If $u\in H^1(\mathbb{R}^{3n})$ is an eigenfunction belonging to an eigenvalue λ in the discrete spectrum of* (1.1)*, then there exists a $\delta>0$ such that*

$$\int e^{2\delta|x|}|u|^2\,dx\,,\ \int e^{2\delta|x|}|\nabla u|^2\,dx\ <\infty\,, \tag{2.6}$$

with $\delta<\sqrt{2d_\lambda}$, where $d_\lambda=\Sigma-\lambda$ and Σ is the infimum of the essential spectrum.

This indicates that particularly for the lowest eigenvalues λ, the assumptions of the following proposition are reasonable.

Proposition 2.3. *If $w^*=e^F u\in H^1(\mathbb{R}^{3n})$, where u is an eigenfunction of a with eigenvalue λ, then w^* solves the adjoint modified problem* (2.5)*.*

Proof. We proceed as above to rewrite $\tilde{a}(\tau,e^F u)$ for $\tau\in\mathcal{D}(\mathbb{R}^{3n})$, using the integration by parts

$$-\frac{1}{2}\int e^F(\nabla F\cdot\nabla\tau)u\,dx = \frac{1}{2}\int e^F(\nabla F\cdot\nabla u)\tau\,dx + \frac{1}{2}\int e^F|\nabla F|^2 u\tau\,dx + \int e^F V_{\mathrm{ee}}u\tau\,dx\,,$$

and that for the compactly supported functions τ, we have $e^F\tau\in H^1(\mathbb{R}^{3n})$. □

Remark 2.4. To give a specific example of what the exponential decay property (2.6) means for w^* from Proposition 2.3, we consider Helium, where $\Sigma=-1$ (cf. [24, XIII.3.A]) and for the ground state we have $\lambda_0<-2.9037$. Since $\sqrt{2d_{\lambda_0}}|x|-\frac{1}{2}|x_1-x_2|\geq\left(\sqrt{2d_{\lambda_0}}-\frac{1}{\sqrt{2}}\right)|x|$, (2.6) still holds with $\delta<\sqrt{2d_{\lambda_0}}-\frac{1}{\sqrt{2}}\approx 1.244$.

3 A Mixed Regularity Estimate

In this section, we study regularity properties relevant for sparse tensor product discretizations of the modified eigenfunctions w from Proposition 2.1. From here on, we restrict the discussion to the two-electron case; in Sections 4 and 5, we put these results in context with computational aspects, where this case already poses formidable difficulties.

We will use certain standard Sobolev spaces of dominating mixed derivatives, employing a notation adapted to our purposes. For $s, k > 0$ and $n \in \mathbb{N}$, we define the Sobolev space $\mathcal{H}^{s,k}(\mathbb{R}^3; n)$ to comprise those $f \in L^2(\mathbb{R}^{3n})$ for which

$$\big\| f \,|\, \mathcal{H}^{s,k}(\mathbb{R}^3; n) \big\|^2 := \int \Big(1 + \prod_i |\xi_i|^{2s}\Big)\Big(1 + \sum_i |\xi_i|^2\Big)^k |\hat{f}|^2 \, d\xi_1 \dots d\xi_n < \infty \,,$$

where $\xi_i \in \mathbb{R}^3$, $i = 1, \dots, n$. In other words, these are the functions for which the mixed derivative of order s, in the meaning of taking s derivatives for each x_i, is in $H^k(\mathbb{R}^{3n})$.

In what follows, we shall use the abbreviations $\|\cdot\|_k$ and $|\cdot|_k$ for norm and seminorm, respectively, on $H^k(\mathbb{R}^{3n})$ and $\|\cdot\|_{s,k}$ for the norm on $\mathcal{H}^{s,k}(\mathbb{R}^3; n)$. To simplify notation for the following consideration of the two-electron case, from here on we will use $x, y \in \mathbb{R}^3$ as single-electron coordinates. Note that for the ground state (i.e. the lowest eigenvalue) of a two-electron system, antisymmetry does not play any role.

Theorem 3.1. *Let $u \in H^1(\mathbb{R}^6)$ solve the electronic Schrödinger equation*

$$-\frac{1}{2}\Delta u + V_{\mathrm{ne}} u + \frac{1}{|x-y|} u = \lambda u \,, \tag{3.1}$$

then $w := \exp(-\frac{1}{2}|x-y|)u$ solves the modified problem

$$Tw := -\frac{1}{2}\Delta w + V_{\mathrm{ne}}\, w - \frac{1}{2}\frac{x-y}{|x-y|} \cdot (\nabla_x - \nabla_y) w - \frac{1}{4} w = \lambda w \tag{3.2}$$

where $w \in \mathcal{H}^{1,1}(\mathbb{R}^3; 2)$.

In the proof of the regularity statement, we will use the following notation: For $\alpha \in I := \{1,2,3\}^2$, let $L_\alpha := \partial_{x_{\alpha_1}} \partial_{y_{\alpha_2}}$. On the spaces $\mathcal{H}^{1,k}(\mathbb{R}^3; 2)$, $k = 0, 1$, we make use of the equivalent norms

$$\|u\|_{1,k} = \Big(\|u\|_k^2 + \sum_{\alpha \in I} \|L_\alpha u\|_k^2 \Big)^{1/2} ,$$

where these spaces can be identified with the closure of $\mathcal{D}(\mathbb{R}^6)$ under the respective norms. We shall also use the space $\mathcal{H}^{1,-1}(\mathbb{R}^3; 2)$ defined by duality and the seminorms

$$|u|_{1,0} = \Big(\sum_{\alpha \in I} \|L_\alpha u\|_0^2 \Big)^{1/2} , \quad |u|_{1,1} = \Big(\sum_{\alpha \in I} |L_\alpha u|_1^2 \Big)^{1/2} .$$

The proof of Theorem 3.1 requires a few preparatory lemmas. We follow the strategy of [29]: First, we show boundedness and ellipticity in $\mathcal{H}^{1,1}$ of the augmented bilinear form

$$b(u, v) := \langle (T + \mu \,\mathrm{I}) u, v \rangle + \sum_{\alpha \in I} \langle L_\alpha (T + \mu \,\mathrm{I}) u, L_\alpha v \rangle \tag{3.3}$$

for $\mu > 0$ large enough; it then remains to be verified that the solution of the corresponding eigenvalue problem, which has the desired smoothness, actually coincides with the solution w of the problem (3.2).

For ellipticity and coercivity, we need appropriate estimates for the potential terms in b. In the case of the single-particle potential V_{ne}, these are the same as in [29].

Lemma 3.2 ([29])**.** *For $u, v \in \mathcal{D}(\mathbb{R}^6)$,*

$$\left| \int V_{\mathrm{ne}} uv \, d(x,y) \right| \lesssim \|u\|_0 |v|_1 \,, \quad \sum_{\alpha \in I} \left| \int L_\alpha (V_{\mathrm{ne}} u) L_\alpha v \, d(x,y) \right| \lesssim |u|_{1,0} |v|_{1,1} \,.$$

For the modified electron-electron operator

$$\tilde{V}_{\mathrm{ee}}(x,y) := -\frac{1}{2}\frac{x-y}{|x-y|}\cdot(\nabla_x - \nabla_y) - \frac{1}{4}\,,$$

a new estimate is required. We make use of the following Hardy-type inequalities, cf. [31, Section 5.4].

Lemma 3.3. *For $v \in \mathcal{D}(\mathbb{R}^6)$,*

$$\int \frac{1}{|x-y|^2} v^2 \, d(x,y) \leq \begin{cases} 4\int |\nabla_x v|^2 \, d(x,y)\,, \\ 4\int |\nabla_y v|^2 \, d(x,y)\,. \end{cases}$$

In contrast to the estimates for the two-electron Coulomb potential in [29], the following result for $\tilde{V}_{\mathrm{ee}}$ for the modified problem does not depend on antisymmetry properties of the involved functions.

Lemma 3.4. *Let $u, v \in \mathcal{D}(\mathbb{R}^6)$, then there exist $C, \tilde{C} > 0$ such that*

$$\left|\int \tilde{V}_{\mathrm{ee}} uv \, d(x,y)\right| \lesssim C|u|_1\|v\|_0 + \frac{1}{4}\|u\|_0\|v\|_0\,,$$

$$\sum_{\alpha\in I}\left|\int L_\alpha(\tilde{V}_{\mathrm{ee}}u)L_\alpha v \, d(x,y)\right| \lesssim \tilde{C}\big(|u|_{1,0}|v|_{1,1} + |u|_{1,1}|v|_{1,0}\big) + \frac{1}{4}|u|_{1,0}|v|_{1,0}\,.$$

Proof. The first estimate is clear. For the second estimate, we need to show

$$\begin{aligned}\sum_{i,j}\left|\int \partial_{x_i}\partial_{y_j}\left(\frac{x-y}{|x-y|}\cdot\nabla_x u\right)\partial_{x_i}\partial_{y_j} v \, d(x,y)\right| \\ \lesssim \left(\sum_{i,j}\|\partial_{x_i}\partial_{y_j}u\|_0^2\right)^{1/2}\left(\sum_{i,j}|\partial_{x_i}\partial_{y_j}v|_1^2\right)^{1/2} \\ + \left(\sum_{i,j}|\partial_{x_i}\partial_{y_j}u|_1^2\right)^{1/2}\left(\sum_{i,j}\|\partial_{x_i}\partial_{y_j}v\|_0^2\right)^{1/2}. \quad (3.4)\end{aligned}$$

The term with $\nabla_y u$ on the left hand side can then be treated analogously.

Using the product rule on the left hand side of (3.4) gives

$$\begin{aligned}\partial_{x_i}\partial_{y_j}\left(\frac{x-y}{|x-y|}\cdot\nabla_x u\right) = \\ \sum_k\left[\left(\partial_{x_i}\partial_{y_j}\frac{x_k-y_k}{|x-y|}\right)\partial_{x_k}u + \left(\partial_{y_j}\frac{x_k-y_k}{|x-y|}\right)\partial_{x_k}\partial_{x_i}u\right. \\ \left. + \left(\partial_{x_i}\frac{x_k-y_k}{|x-y|}\right)\partial_{x_k}\partial_{y_j}u + \frac{x_k-y_k}{|x-y|}\big(\partial_{x_k}\partial_{x_i}\partial_{y_j}u\big)\right]. \quad (3.5)\end{aligned}$$

The integrals for the rightmost term with third derivatives of u can be estimated directly by $|u|_{1,1}|v|_{1,0}$, since the coefficient is uniformly bounded.

We next turn to the term in (3.5) containing second derivatives of the coefficient,

$$\begin{aligned}
&\sum_{i,j}\left|\int\left(\partial_{x_i}\partial_{y_j}\frac{x-y}{|x-y|}\right)\cdot(\nabla_x u)\,(\partial_{x_i}\partial_{y_j}v)\,d(x,y)\right| \\
&\quad\lesssim\int\frac{1}{|x-y|^2}|\nabla_x u|\Big(\sum_{i,j}\big|\partial_{x_i}\partial_{y_j}v\big|^2\Big)^{1/2}d(x,y) \\
&\quad\lesssim\left(\int\frac{1}{|x-y|^2}|\nabla_x u|^2\,d(x,y)\right)^{1/2}\left(\sum_{i,j}\int\frac{1}{|x-y|^2}\big|\partial_{x_i}\partial_{y_j}v\big|^2\,d(x,y)\right)^{1/2},
\end{aligned}$$

and using Lemma 3.3,

$$\lesssim\left(\sum_{j,k}\int\big|\partial_{y_j}\partial_{x_k}u\big|^2\,d(x,y)\right)^{1/2}\left(\sum_{k,i,j}\int\big|\partial_{x_k}\partial_{x_i}\partial_{y_j}v\big|^2\,d(x,y)\right)^{1/2}=|u|_{1,0}|v|_{1,1}\,.$$

For the middle terms in (3.5) we obtain, again using Lemma 3.3,

$$\begin{aligned}
&\sum_{i,j}\left|\int\left(\partial_{y_j}\frac{x-y}{|x-y|}\right)\cdot(\nabla_x\partial_{x_i}u)\,(\partial_{x_i}\partial_{y_j}v)\,d(x,y)\right| \\
&\qquad\lesssim\int\frac{1}{|x-y|}\Big(\sum_{i,k}\big|\partial_{x_i}\partial_{x_k}u\big|^2\Big)^{1/2}\Big(\sum_{i,j}\big|\partial_{x_i}\partial_{y_j}v\big|^2\Big)^{1/2}d(x,y) \\
&\qquad\lesssim\left(\sum_{i,k}\int\frac{1}{|x-y|^2}\big|\partial_{x_i}\partial_{x_k}u\big|^2\,d(x,y)\right)^{1/2}\left(\sum_{i,j}\int\big|\partial_{x_i}\partial_{y_j}v\big|^2\,d(x,y)\right)^{1/2} \\
&\qquad\lesssim\left(\sum_{i,j,k}\int\big|\partial_{x_k}\partial_{x_i}\partial_{y_j}u\big|^2\,d(x,y)\right)^{1/2}\left(\sum_{i,j}\int\big|\partial_{x_i}\partial_{y_j}v\big|^2\,d(x,y)\right)^{1/2}=|u|_{1,1}|v|_{1,0}\,,
\end{aligned}$$

where the roles of x_i and y_j can be interchanged. Altogether, this shows (3.4). □

Proof of Theorem 3.1. Boundedness of b as in (3.3) on $\mathcal{H}^{1,1}(\mathbb{R}^3;2)$ follows with Lemmas 3.2 and 3.4. Furthermore, with constants $c,\tilde c>0$, for $v\in\mathcal{H}^{1,1}(\mathbb{R}^3;2)$ we have

$$b(v,v)\geq\frac12\|v\|_1^2-c\|v\|_0|v|_1+\Big(\mu-\frac14\Big)\|v\|_0^2+\frac12|v|_{1,1}^2-\tilde c|v|_{1,0}|v|_{1,1}+\Big(\mu-\frac14\Big)|v|_{1,0}^2\,,$$

hence ellipticity on $\mathcal{H}^{1,1}(\mathbb{R}^3;2)$ follows if μ is chosen large enough.

Let $w\in H^1(\mathbb{R}^6)$ be a solution of $(T+\mu\,\mathrm{I})w=(\lambda+\mu)w$. Then $L_\alpha w\in H^{-1}(\mathbb{R}^6)$ for $\alpha\in I$, i.e. $w\in\mathcal{H}^{1,-1}(\mathbb{R}^3;2)$, and the variational problem

$$b(\psi,v)=(\lambda+\mu)\Big[\langle w,v\rangle+\sum_\alpha\langle L_\alpha w,L_\alpha v\rangle\Big]\,,\quad v\in\mathcal{H}^{1,1}(\mathbb{R}^3;2)$$

has a unique solution $\psi\in\mathcal{H}^{1,1}(\mathbb{R}^3;2)$.

Let $v\in\mathcal{S}(\mathbb{R}^6)$. Then we can integrate by parts on the right hand side to obtain

$$\langle w,v\rangle+\sum_\alpha\langle L_\alpha w,L_\alpha v\rangle=\langle w,v+\Delta_x\Delta_y v\rangle\,.$$

It can be verified as in [29] that one can integrate by parts in b as well, which yields

$$b(\psi,v)=\langle(T+\mu\,\mathrm{I})\psi,v+\Delta_x\Delta_y v\rangle\,.$$

Hence for any $v \in \mathcal{S}(\mathbb{R}^6)$,

$$\langle (T + \mu\,\mathrm{I})\psi, v + \Delta_x\Delta_y v\rangle = (\lambda + \mu)\langle w, v + \Delta_x\Delta_y v\rangle\,.$$

For any given $\varphi \in \mathcal{S}(\mathbb{R}^6)$, we can find $v \in \mathcal{S}(\mathbb{R}^6)$ such that $v + \Delta_x\Delta_y v = \varphi$. Consequently, $\langle (T + \mu\,\mathrm{I})\psi, \varphi\rangle = (\lambda + \mu)\langle w, \varphi\rangle$ for any $\varphi \in \mathcal{S}(\mathbb{R}^6)$, by density also for any $\varphi \in H^1(\mathbb{R}^6)$. By ellipticity, this implies $\psi = w$. □

Remark 3.5. One can modify the proof along the lines of [31, Chapter 5.1] to additionally obtain exponential decay of mixed derivatives, which in the present two-electron case can be phrased as follows: there exists $\bar\gamma > 0$ such that for any γ, $0 < \gamma < \bar\gamma$,

$$\exp\big(\gamma(|x| + |y|)\big) w \in \mathcal{H}^{1,1}(\mathbb{R}^3; 2)\,. \tag{3.6}$$

Remark 3.6. For more than two electrons, the modified variational formulation has been shown to be satisfied in Proposition 2.1 as well. The proof of a corresponding regularity estimate, however, leads to further complications due to the higher-order singularities arising in the additional three-electron integrals; since we will focus on the two-electron case in what follows, we do not pursue this further at this point.

Remark 3.7. As to be expected in view of Proposition 2.3, the proof of Theorem 3.1 does not carry over to the adjoint problem (2.5). The point where it breaks down is rather subtle: In order to be able to carry out the necessary integration by parts for moving all higher derivatives onto the test function v, the nonsymmetric terms in the augmented bilinear form need to be taken slightly differently as

$$\sum_\alpha \Big\langle L_\alpha\Big(\frac{x-y}{|x-y|}u\Big), L_\alpha(\nabla_x - \nabla_y)v\Big\rangle\,.$$

This means that in the proof of an analogue of Lemma 3.4, after using the product rule we need to estimate integrals of the form

$$\int \frac{1}{|x-y|^2}|u|\,|(\nabla_x - \nabla_y)\partial_{x_i}\partial_{y_j}v|\,d(x,y)\,.$$

Now we cannot split the singular factor and estimate part of the singularity with v as before, since this would leave the resulting derivatives on u and v unbalanced. Proceeding instead directly via the Cauchy-Schwarz inequality, the required estimate

$$\int \frac{1}{|x-y|^4}|u|^2\,d(x,y) \lesssim \sum_{i,j}\int |\partial_{x_i}\partial_{y_j}u|^2\,d(x,y)$$

holds true only if u vanishes on $\{x = y\}$, see [29]. Thus following the reasoning there gives only $\mathcal{H}^{1/2,1}(\mathbb{R}^3; 2)$ for the adjoint problem, i.e. the same regularity as for the eigenfunctions of the unmodified bilinear form a.

4 Approximation by Wavelets

In this section, we discuss the implications of the regularity estimate of Theorem 3.1 for hyperbolic wavelet approximation of eigenfunctions of the correlated formulation (3.2), as well as some basic issues concerning numerical realization.

The following will be based on an orthonormal wavelet basis for $L^2(\mathbb{R})$ with scaling function φ and wavelet ψ, where we set

$$\psi_{j,k,0} = 2^{j/2}\varphi(2^j\cdot - k)\,,\quad \psi_{j,k,1} = 2^{j/2}\psi(2^j\cdot - k)\,,$$

and we use the basis starting (without restriction of generality) from scaling functions at level $j = 0$,

$$\{\psi_\lambda\}_{\lambda\in\bar\Lambda^{(1)}}, \quad \bar\Lambda^{(1)} := \{(0,k,0)\colon k\in\mathbb{Z}\}\cup\{(j,k,1)\colon j\geq 0, k\in\mathbb{Z}\}\,.$$

Further details on the choice of wavelets and the reasons for starting from an orthonormal basis are given in Subsection 4.2. We further introduce the abbreviations $|\lambda| := j$, $s(\lambda) := s$ for $\lambda = (j,k,s) \in \bar\Lambda^{(1)}$. Whenever the basis functions are sufficiently regular, $\{2^{-k|\lambda|}\psi_\lambda\}_{\lambda\in\bar\Lambda^{(1)}}$ is a Riesz basis for $H^k(\mathbb{R})$, $k = 1, 2$, which we assume for what follows. This regularity assumption could be weakened at the expense of additional technical difficulties, which we avoid here to keep the exposition simple.

For higher dimensions d, let $\gamma_{(j,k,s)} := \bigotimes_{i=1}^d \psi_{j,k_i,s_i}$ and

$$\bar\Lambda^{(d)} := \big\{(0,k,0)\colon k\in\mathbb{Z}^d\big\}\cup\big\{(j,k,s)\colon j\geq 0, k\in\mathbb{Z}^d, s\in\{0,1\}^d\setminus\{0\}\big\}\,,$$

then $\{\gamma_\lambda\}_{\lambda\in\bar\Lambda^{(d)}}$ is an orthonormal basis of $L^2(\mathbb{R}^d)$ and $\{2^{-k|\lambda|}\gamma_\lambda\}_{\lambda\in\bar\Lambda^{(d)}}$ is a Riesz basis of $H^k(\mathbb{R}^d)$, $k = 1, 2$.

4.1 Hyperbolic wavelet bases

For the n-electron case, we are interested in subsets of the tensor product basis $(\bar\Lambda^{(3)})^n$,

$$\Psi_\lambda = \bigotimes_{i=1}^n \gamma_{\lambda_i}\,, \quad \lambda = (\lambda_1,\ldots,\lambda_n)\,,$$

where Riesz bases for $H^1(\mathbb{R}^{3n})$ and $\mathcal{H}^{s,1}(\mathbb{R}^3;n)$, $0 < s \leq 1$, are given by

$$\Big\{(2^{2|\lambda_1|}+\ldots+2^{2|\lambda_n|})^{-1/2}\Psi_\lambda\Big\} \quad\text{and}\quad \Big\{(2^{2|\lambda_1|}+\ldots+2^{2|\lambda_n|})^{-1/2}\,2^{-s(|\lambda_1|+\ldots+|\lambda_n|)}\Psi_\lambda\Big\}\,,$$

respectively, see e.g. [10], and we have the norm equivalences

$$\|u\|_1 \sim \sum_{\lambda\in\bar\Lambda^{(n)}} (2^{2|\lambda_1|}+\ldots+2^{2|\lambda_n|})|\langle u,\Psi_\lambda\rangle|^2\,, \tag{4.1a}$$

$$\|u\|_{s,1} \sim \sum_{\lambda\in\bar\Lambda^{(n)}} (2^{2|\lambda_1|}+\ldots+2^{2|\lambda_n|})2^{2s(|\lambda_1|+\ldots+|\lambda_n|)}|\langle u,\Psi_\lambda\rangle|^2\,. \tag{4.1b}$$

In what follows, for $u\in L^2(\mathbb{R}^{3n})$ and $\Lambda\subset(\bar\Lambda^{(3)})^n$,

$$u_\Lambda := \sum_{\lambda\in\Lambda}\langle u,\Psi_\lambda\rangle\Psi_\lambda\,.$$

A family of hyperbolic wavelet bases with discretization parameter $L\in\mathbb{N}_0$ is given by the index sets

$$\Lambda_L^{(n)} := \Big\{\lambda\in(\bar\Lambda^{(3)})^n\colon |\lambda_1|+\ldots+|\lambda_n|\leq L\Big\}\,.$$

In a finite element context, this construction is known as a sparse grid.

For this type of discretization, the regularity estimate of Theorem 3.1 and the exponential decay of Theorem 2.2 can be combined to a simple approximation result for eigenfunctions of (3.2), which is certainly not the best possible, but is included here rather for illustrative purposes: We restrict the approximation to a subset Λ of indices from the hyperbolic wavelet basis $\Lambda_L^{(2)}$ that is confined to a region around the origin, and estimate separately the error due to truncation in space by Theorem 2.2 and the error due to truncation in level L by Theorem 3.1.

For eigenfunctions $w\in\mathcal{H}^{1,1}(\mathbb{R}^3;2)$ as in Theorem 3.1, which also satisfy (2.6), the following applies with $s = 1$.

Theorem 4.1. *If $u \in \mathcal{H}^{s,1}(\mathbb{R}^3;2)$, $0 < s \le 1$, satisfies a decay condition (2.6) with some $\delta > 0$, then there exists a $C > 0$ depending on δ and ψ such that for each L there exists a subset $\Lambda \subset \Lambda_L^{(2)}$ with*

$$\inf_{v \in \operatorname{span}\{\Psi_\lambda :\, \lambda \in \Lambda\}} \|u - v\|_1 \le C 2^{-sL} \|u\|_{s,1}$$

and $|\Lambda| \sim L^7 2^{3L}$.

Proof. Fix $M > 0$. For $R > M$, let $\eta_R \in C^\infty(\mathbb{R}^6)$, $0 \le \eta_R \le 1$, such that $\eta_R \equiv 0$ on $B_{R-M}(0)$, $\eta_R \equiv 1$ on $\mathbb{R}^6 \setminus B_R(0)$ and $\|\nabla \eta_R\|_\infty$ is bounded uniformly in R. Then with a $c > 0$ independent of R and $v \in H^1(\mathbb{R}^6)$,

$$\|\eta_R v\|_1 \le c \|v \mid H^1(\mathbb{R}^6 \setminus B_{R-M}(0))\| \,.$$

Let $\Lambda_R = \{\lambda \in (\bar\Lambda^{(3)})^2 :\, \lambda \in \Lambda_L^{(2)},\, \operatorname{supp} \Psi_\lambda \cap B_R(0) \ne \emptyset\}$. Then

$$u - u_{\Lambda_R} = \sum_{\lambda \notin \Lambda_L^{(2)}} \langle u, \Psi_\lambda \rangle \Psi_\lambda + \sum_{\lambda \in \Lambda_L^{(2)} \wedge \lambda \notin \Lambda_R} \langle u, \Psi_\lambda \rangle \Psi_\lambda \,.$$

Using the wavelet characterization (4.1a),

$$\|u - u_{\Lambda_R}\|_1^2 \lesssim \sum_{\lambda \notin \Lambda_L^{(2)}} (2^{2|\lambda_1|} + 2^{2|\lambda_2|}) |\langle u, \Psi_\lambda \rangle|^2 + \sum_{\lambda \notin \Lambda_R} (2^{2|\lambda_1|} + 2^{2|\lambda_2|}) |\langle u, \Psi_\lambda \rangle|^2 \,.$$

Now on the one hand, by (4.1b)

$$\begin{aligned} \sum_{\lambda \notin \Lambda_L^{(2)}} (2^{2|\lambda_1|} + 2^{2|\lambda_2|}) |\langle u, \Psi_\lambda \rangle|^2 &\le 2^{-2s(L+1)} \sum_{|\lambda_1| + |\lambda_2| > L} (2^{2|\lambda_1|} + 2^{2|\lambda_2|}) 2^{2s(|\lambda_1| + |\lambda_2|)} |\langle u, \Psi_\lambda \rangle|^2 \\ &\lesssim 2^{-2sL} \|u\|_{s,1}^2 \,, \end{aligned}$$

on the other hand, since $u = \eta_R u$ outside $B_R(0)$,

$$\begin{aligned} \sum_{\lambda \notin \Lambda_R} (2^{2|\lambda_1|} + 2^{2|\lambda_2|}) |\langle u, \Psi_\lambda \rangle|^2 &\le \sum_\lambda (2^{2|\lambda_1|} + 2^{2|\lambda_2|}) |\langle \eta_R u, \Psi_\lambda \rangle|^2 \lesssim \|\eta_R u\|_1^2 \\ &\le c^2 e^{-2\delta(R-M)} \int e^{2\delta|(x,y)|} \left(|u|^2 + |\nabla u|^2\right) d(x,y) \lesssim e^{-2\delta R} \,. \end{aligned}$$

In summary, $\|u - u_{\Lambda_R}\|_1 \lesssim e^{-\delta R} + 2^{-sL} \|u\|_{s,1}$, where we choose $R \sim L$ to balance the two expressions on the right hand side. The number of $\lambda \in \Lambda := \Lambda_R$ with $|\lambda_1| = j_1$, $|\lambda_2| = j_2$ is then proportional to $(L 2^{j_1})^3 (L 2^{j_2})^3$, and by

$$|\Lambda| \sim L^6 \sum_{0 \le j_1 + j_2 \le L} 2^{3(j_1 + j_2)} = L^6 \sum_{l=0}^{L} (l+1) 2^{3l}$$

we obtain $|\Lambda| \sim L^7 2^{3L}$. □

Remark 4.2. Although a deeper analysis will allow to obtain better convergence in terms of $|\Lambda|$, the above already gives almost the basic complexity in L of a three-dimensional problem discretized by uniformly refined piecewise linear basis functions. For comparison, we can apply Theorem 4.1 to a direct hyperbolic wavelet discretization of the standard formulation (3.1): It has been shown in [29] that the corresponding eigenfunctions are in $\mathcal{H}^{1/2,1}$, and one obtains a convergence rate of $2^{-L/2}$. In terms of our previous analogy, in this case the complexity corresponds to the discretization of of a six-dimensional problem by uniformly refined piecewise linear basis functions; hence in the two-electron case, it is only in the explicitly correlated formulation that the regularity results guarantee an advantage of hyperbolic wavelets over a full tensor product. It needs to be noted, however, that for the convergence of eigenvalues in the nonsymmetric modified problem we can only expect a rate of $2^{-3L/2}$ due to the slower convergence of adjoint eigenfunctions.

At this point, we do not pursue the matter of sharper approximation error estimates further, but discuss suitable choices of wavelets before turning to further issues of practical realization of such approximations.

4.2 Suitable Wavelet Bases

For obtaining approximations as in Theorem 4.1 as solutions of wavelet-Galerkin discretized eigenvalue problems, orthonormality of the underlying basis becomes almost a necessity, in particular when keeping the many-electron case in mind.

First, given orthonormality, the condition numbers of the involved matrices, i.e. of the discretizations of identity, Laplacian, and potential operators remain uniformly bounded with respect to the number of electrons n, whereas they would increase exponentially in n for any non-orthonormal basis. A second important point is that orthogonality allows efficient application of matrices arising from anisotropic tensor products in the higher-dimensional case. For bases that do not have at least some partial orthogonality properties, no algorithm of comparable scaling appears to be known. A further consequence of orthonormality is that the discretization of the identity is the identity matrix, hence instead of a generalized eigenvalue problem only a standard eigenvalue problem needs to be solved. Finally, in view of extensibility to several electrons, antisymmetry properties are quite difficult to enforce unless orthonormal basis functions are used.

Using a basis that is also stable, a simple diagonal scaling provides asymptotically optimal preconditioning for the discretized Hamiltonian; wavelet bases satisfy this requirement if they possess sufficient regularity and high enough vanishing moments. For orthonormal wavelet bases, the number of vanishing moments also determines the order of polynomial reproduction, hence it is desirable to have the option to use a higher number of vanishing moments for higher-order approximation.

From a computational point of view, basis functions should be compactly supported. This facilitates an efficient computation of integrals, greatly simplifies applying discretization matrices efficiently to a vector, and allows for local refinement of the subset of the basis that is used in the computation.

Altogether, considering these rather restrictive requirements, the least asymmetric Daubechies wavelets [9] appear to be a suitable choice; a possible alternative would be the piecewise polynomial continuous multiwavelets of [11] as used for instance in [10].

Note, however, that for the nonsymmetric formulation (2.4) one could also consider using a pair of different, biorthogonal bases for ansatz and test space in a Petrov-Galerkin-type discretization. To what extent this can be advantageous remains to be investigated.

5 Exploiting Structural Properties for Applying Operators

Up to now, we were dealing with the question of how well eigenfunctions in an explicitly correlated formulation can in principle be approximated by hyperbolic wavelets. In this section we turn to basic prerequisites for the actual computation of such approximations as solutions of discretized eigenvalue problems.

We consider a wavelet-Galerkin discretization of the modified eigenvalue problem (2.4) defined by a subset of wavelet indices $\Lambda \subset \Lambda_L^{(2)}$, where under our assumption of an orthonormal basis, the discretized problem for the eigenfunction coefficient vector (u_μ) becomes

$$\sum_{\mu\in\Lambda} \tilde{a}(\Psi_\mu, \Psi_\nu) u_\mu = \lambda u_\nu \quad \text{for } \nu \in \Lambda\,. \tag{5.1}$$

Such an approach with Λ fixed a priori corresponds to the regularity and approximation results from the previous sections, but the concepts discussed in the following may also be useful in the framework of adaptive wavelet schemes.

We again focus on the basic case of two electrons, for simplicity restricting our discussion to the atomic case corresponding to Helium as in (1.2). Recall that $\mu = (\mu_1, \mu_2)$, $\nu = (\nu_1, \nu_2)$ and $\Psi_\mu = \gamma_{\mu_1} \otimes \gamma_{\mu_2}$, $\Psi_\nu = \gamma_{\nu_1} \otimes \gamma_{\nu_2}$; the matrix entries read

$$\tilde{a}(\Psi_\mu, \Psi_\nu) = h_{\mu_1\nu_1}\delta_{\mu_2\nu_2} + \delta_{\mu_1\nu_1}h_{\mu_2\nu_2} - \frac{1}{2}\int_{\mathbb{R}^6} \frac{x-y}{|x-y|} \cdot (\nabla_x - \nabla_y)\Psi_\mu\Psi_\nu \, d(x,y) - \frac{1}{4}\delta_{\mu\nu} \tag{5.2}$$

with the single-electron matrices

$$h_{\lambda\kappa} = \frac{1}{2}\int_{\mathbb{R}^3} \nabla\gamma_\lambda \cdot \nabla\gamma_\kappa \, dx - \int_{\mathbb{R}^3} \frac{2}{|x|}\gamma_\lambda\gamma_\kappa \, dx\,, \quad \lambda, \kappa \in \bar{\Lambda}^{(3)}\,. \tag{5.3}$$

In general, Galerkin discretization matrices arising from hyperbolic wavelet bases are almost dense, the most problematic aspect being the high number of nonzero matrix entries with $|\mu_1|, |\nu_2|$ small and $|\mu_2|, |\nu_1|$ large or vice versa. Therefore, assembling the matrix in (5.1) explicitly is out of the question.

Instead, we aim to construct algorithms for applying this matrix with reasonable complexity. Consequently, we are interested in iterative solvers that only require the evaluation of matrix-vector products, and since our objective is to compute a few lowest eigenvalues and an effective diagonal preconditioner is available in our setting, slight refinements of basic inverse iteration-type algorithms are sufficient.

The basic strategy we follow to achieve fast application of matrices is to exploit tensor product structures. If the underlying operators have such structure themselves, this can be used directly as in the well-known basic algorithm outlined in the next subsection. Especially in view of the two-electron operator, such a scheme by itself is not sufficient for (5.2). One could proceed by directly using matrix compression based on the vanishing moment properties of wavelets, resulting in a perturbed application of operators. However, such a method is still highly expensive in this setting. This leads us to consider approximations for the coefficients that introduce auxiliary tensor product structures, which can also be seen as a very problem-specific type of compression, in Section 5.3.

5.1 Matrix-Vector Products for Hyperbolic Wavelet Discretizations

For operators that have a tensor product structure, an efficient scheme for applying discretization matrices arising from subsets of $\Lambda_L^{(n)}$ can be obtained by exploiting this structure for certain blocks corresponding to different combinations of levels. Consider, for instance, a tensor product operator $A = A^{(1)} \otimes A^{(2)}$ with $A^{(k)}\colon H^1(\mathbb{R}^3) \to H^{-1}(\mathbb{R}^3)$, $k = 1, 2$, and its wavelet-Galerkin discretization for $\Lambda \subset \Lambda_L^{(2)}$,

$$A_\Lambda = \big(a_{\nu\mu}\big)_{\nu,\mu\in\Lambda}\,, \quad a_{\nu\mu} = \langle A\Psi_\mu, \Psi_\nu\rangle = \langle A^{(1)}\gamma_{\mu_1}, \gamma_{\nu_1}\rangle\langle A^{(2)}\gamma_{\mu_2}, \gamma_{\nu_2}\rangle\,. \tag{5.4}$$

In the scheme described in [25] for applying A_Λ to a vector, the matrix is decomposed into blocks corresponding to fixed levels. To this end, we define subsets of indices

$$\Lambda_{(j_1,j_2),(s_1,s_2)} = \{\lambda \in \Lambda\colon |\lambda_1| = j_1,\, |\lambda_2| = j_2,\, s(\lambda_1) = s_1,\, s(\lambda_2) = s_2\}\,,$$

suppressing in the following for better clarity the subscripts s_1, s_2 describing the combination of one-dimensional wavelets and scaling functions. If these subsets of indices have a tensor product structure, then each block $A_{(i_1,i_2),(j_1,j_2)} = (a_{\nu\mu})_{\nu\in\Lambda_{(i_1,i_2)},\mu\in\Lambda_{(j_1,j_2)}}$ can be written as a Kronecker product of corresponding blocks $A^{(k)}_{i_k,j_k}$ of the discretization of $A^{(k)}$, and these in turn as matrix products

$$A_{(i_1,i_2),(j_1,j_2)} = A^{(1)}_{i_1,j_1} \otimes A^{(2)}_{i_2,j_2} = \begin{cases} (A^{(1)}_{i_1,j_1} \otimes \mathrm{I}_{i_2,j_2}) \cdot (\mathrm{I}_{i_1,j_1} \otimes A^{(2)}_{i_2,j_2}) & \text{(a)}\,, \\ (\mathrm{I}_{i_1,j_1} \otimes A^{(2)}_{i_2,j_2}) \cdot (A^{(1)}_{i_1,j_1} \otimes \mathrm{I}_{i_2,j_2}) & \text{(b)}\,, \end{cases} \tag{5.5}$$

where I_{i_k,j_k} denote the identities on the index subsets. Now for each block, one can pick the order of applying the two factors that minimizes the size of the resulting intermediates: case (a) is used if $j_1 + i_2 \le i_1 + j_2$, case (b) otherwise.

It can be shown also for higher dimension (cf. [25]) that by this strategy, tensor product operators discretized by hyperbolic wavelets can be applied with a complexity that is, up to dimension-dependent powers of L, of the same order as the number of unknowns. Furthermore, the scheme can also be applied to more general subsets Λ of the full tensor product basis $(\bar{\Lambda}^{(3)})^2$.

Note that when using an orthonormal basis, the procedure simplifies further if some of the lower-dimensional factors $A^{(k)}$ are identities, as in case of the one-electron parts in (5.2). Thus the basic technique of (5.5) by itself is already useful for the single-electron terms in (5.2), though not for the two-electron operator.

5.2 Evaluation of Integrals

The computation of matrix entries for Galerkin discretizations based on Daubechies or similar wavelets is complicated by the lack of a closed-form representation and the inefficiency of direct quadrature. However, certain basic integrals of wavelets can be reduced by the scaling relations to integrals of scaling functions on the lowest level, which in turn can be obtained from a constrained eigenvalue problem that involves only the scaling coefficients (cf. [2, 8]); in particular, this suffices for the computation of certain basic building blocks that appear in the matrix entries of (5.2), for instance

$$\int_{\mathbb{R}} \psi_{j_1,k_1,s_1}\psi_{j_2,k_2,s_2}\,dx\,, \quad \int_{\mathbb{R}} \psi''_{j_1,k_1,s_1}\psi_{j_2,k_2,s_2}\,dx\,, \quad \int_{\mathbb{R}} \psi_{j_0,k_0,s_0}\psi_{j_1,k_1,s_1}\psi_{j_2,k_2,s_2}\,dx\,.$$

This is sufficient for the Laplacian terms in (5.2). As pointed out in [13], the procedure can, in principle, be adapted to the computation of certain integrals for the potential terms, e.g.

$$\int_{\mathbb{R}^3} \frac{1}{|x|}\gamma_{(j,k,s)}\,dx\,, \quad \int_{\mathbb{R}^6} \frac{x-y}{|x-y|}\cdot\big(\nabla_x\gamma_{(j,k_1,s_1)}(x)\big)\gamma_{(j,k_2,s_2)}(y)\,d(x,y)\,,$$

where integrals appearing in (5.2) can be obtained for example by

$$\int \frac{1}{|x|}\gamma_\mu\gamma_\nu\,dx \approx \sum_{\lambda\in\Lambda\subset\bar{\Lambda}^{(3)}} \int \frac{1}{|x|}\gamma_\lambda\,dx \int \gamma_\lambda\gamma_\mu\gamma_\nu\,dx\,. \tag{5.6}$$

Unfortunately, even when using a sparse approximation of the potential the approach becomes very expensive in practice, which is mostly due to the fairly large support of basis functions, rendering even storing the full discretization matrix on the left hand side of (5.6) hardly feasible; the problem is of course even more pronounced for the two-electron integrals. Hence we do not use this direct approach for the potential operators, but introduce an additional tensor product structure by using a further approximation, as discussed in the following subsection.

5.3 Separable Approximation of Potentials

The basic idea for reducing the complexity of applying one- and two-electron potential operators in (5.2) is to replace them by separable approximations, in the present two-electron case[2]

$$\frac{1}{|x|} \approx \sum_k f_k(x_1)f_k(x_2)f_k(x_3)\,, \quad \frac{x-y}{|x-y|} \approx (x-y)\sum_k g_k(x_1-y_1)g_k(x_2-y_2)g_k(x_3-y_3)\,. \tag{5.7}$$

[2]Here it is possible to choose $g_k = f_k$, although a different choice, e.g. based on a weighted approximation estimate as in Section 5.3.2, can be advantageous.

In these particular cases, surprisingly efficient approximations can be obtained when f_k, g_k are chosen as Gaussian functions $\omega_k^{1/3} \exp(-\alpha_k |\cdot|^2)$ with suitable $\omega_k, \alpha_k > 0$. The existence of such approximations is closely related to the integral representation

$$\frac{1}{|x|} = \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-|x|^2 t^2} \, dt$$

that forms the basis of efficient integral evaluation for Gaussian-type orbitals.

In connection with wavelet bases, in [21, 6, 12] an approach for efficient computation of individual integrals using separable approximations has been developed, focusing on the efficient computation of individual integrals and wavelet coefficients and accordingly tailored error estimates. Since for our purposes we want to avoid explicitly assembling matrices whenever possible, we are rather interested in a different point of view: We replace the potential terms in (5.2) by separable approximations (5.7) and exploit the tensor product structure for applying the operators efficiently.

To quantify the error in computed eigenvalues ultimately caused by such an approximation when combined with the wavelet discretization, we derive estimates for the error in the potential terms in operator norm; to the author's knowledge, such estimates have not appeared in the literature so far. Before turning to the error analysis, let us briefly discuss the algorithmic use of expansions of the form (5.7).

As a prerequisite, we need a method for integrating products of Gaussians with wavelets. This can be done using an auxiliary wavelet family where the scaling function has vanishing moments, i.e. satisfies $\int x^k \varphi \, dx = 0$ for some range $k = 1, \dots, K$, for example the Deslaurier-Dubuc-Sweldens wavelets [26] as suggested in [13].

As in Subsection 5.1, to make use of the tensor product structure we consider blocks of the discretization matrix belonging to fixed levels and wavelet subspaces, and assume the index set Λ defining the discretization to have a tensor structure on these blocks as well. For the one-electron potentials, it is straightforward to arrive at a scheme that consists in applying matrices that operate along single dimensions, hence we only describe the more complicated scheme for the two-electron potential in what follows. We need to apply a Kronecker product of three matrices, each coupling two dimensions, of the form

$$T_{k_1,k_2,l_1,l_2} = \int_{\mathbb{R}^2} e^{-\alpha(x-y)^2} \psi_{(j_1,k_1,s_1)}(x) \, \psi_{(j_2,k_2,s_2)}(y) \, \psi_{(i_1,l_1,t_1)}(x) \, \psi_{(i_2,l_2,t_2)}(y) \, d(x,y)$$

with (j_1, j_2), (i_1, i_2) and (s_1, s_2), (t_1, t_2) being fixed. Due to anisotropy of the basis functions, for certain combinations of levels j_1, j_2, i_1, i_2, these matrices are dense.

Depending on the combination of exponents α and the wavelet levels, many of the entries will actually be negligible due to vanishing moment properties of the wavelets; an example of an extreme case would be small α, large j_1, j_2, and small i_1, i_2. In such situations, matrix compression techniques yield an efficient approximate application. In combinations where such methods are not effective, an alternative approach along the lines of (5.5) is possible: We construct, potentially using some suitable second wavelet basis, a sparse approximation of the diagonal coefficient,

$$e^{-\alpha(x-y)^2} \approx \sum_{(\ell,q,r) \in \Delta \subset \bar{\Lambda}^{(2)}} c_{\ell,q_1-q_2,r} \, \psi_{\ell,q_1,r_1}(x) \, \psi_{\ell,q_2,r_2}(y) \,, \tag{5.8}$$

which will converge rapidly with respect to ℓ if α is small, and will be concentrated along the diagonal if α is large.

To approximate $v_{l_1,l_2} = \sum_{k_1,k_2} T_{k_1,k_2,l_1,l_2} u_{k_1,k_2}$ for a vector u, if $j_1 + i_2 \le i_1 + j_2$ we proceed as follows:

1. Precompute $m_{\ell,q_1,r_1,k_2,l_2} := \sum_{q_2,r_2} c_{\ell,q_1-q_2,(r_1,r_2)} \int_{\mathbb{R}} \psi_{\ell,q_2,r_2} \, \psi_{j_2,k_2,s_2} \, \psi_{i_2,l_2,t_2} \, dx$,

2. Compute the intermediate quantities $w_{\ell,q_1,r_1,k_1,l_2} := \sum_{k_2} m_{\ell,q_1,r_1,k_2,l_2}\, u_{k_1,k_2}$,

3. Compute the result $v_{l_1,l_2} := \sum_{k_1} \sum_{\ell,q_1,r_1} w_{\ell,q_1,r_1,k_1,l_2} \int_{\mathbb{R}} \psi_{\ell,q_1,r_1}\, \psi_{j_1,k_1,s_1}\, \psi_{i_1,l_1,t_1}\, dx$.

If $j_1 + i_2 > i_1 + j_2$, the roles of the dimensions are reversed. One can gain some more efficiency by applying the above algorithm in the range of those ℓ for which this gives an advantage, and again apply the matrix corresponding to the remainder of the sum over ℓ in (5.8) approximately using compression strategies.

Note that one need not use the same separable approximation for every block of the discretization matrix, but for simplicity we restrict the following first analysis to this case.

5.3.1 Approximation of One-Electron Potentials

For the one-electron Coulomb potentials, we consider an exponential approximation that satisfies, cf. [5, Section 4.3],

$$\Big\| \frac{1}{\sqrt{t}} - \sum_{k=1}^{N} \omega_k \exp(-\alpha_k t) \Big\|_{\infty,[1,\infty]} \leq 8\sqrt{2} \exp(-\pi\sqrt{N/2}) =: \delta(N)\,,$$

which leads to the following error estimate for the approximation of the Coulomb potential on $\mathbb{R}^3 \setminus B_r(0)$,

$$\Big\| \frac{1}{|x|} - \sum_{k=1}^{N} \frac{\omega_k}{r} \exp\Big(-\frac{\alpha_k}{r^2}|x|^2\Big) \Big\|_{\infty,\{|x|\geq r\}} \leq r^{-1}\delta(N)\,. \tag{5.9}$$

Theorem 5.1. *For any $\Lambda \subseteq \bar{\Lambda}^{(3)}$, let $A_\Lambda, A_\Lambda^{(N)} \colon H^1(\mathbb{R}^3) \to H^{-1}(\mathbb{R}^3)$ be the operators defined by the matrices $(a_{\nu\mu})_{\nu,\mu\in\Lambda}$, $(a_{\nu\mu}^{(N)})_{\nu,\mu\in\Lambda}$ with*

$$a_{\nu\mu} = \int_{\mathbb{R}^3} \frac{1}{|x|} \gamma_\mu \gamma_\nu \, dx\,, \quad a_{\nu\mu}^{(N)} = \sum_{k=1}^{N} \frac{\omega_k}{r} \int_{\mathbb{R}^3} \exp\Big(-\frac{\alpha_k}{r^2}|x|^2\Big) \gamma_\mu \gamma_\nu \, dx\,,$$

where $\{\omega_k\}$, $\{\alpha_k\}$ satisfy (5.9), then for any $\beta \in (0,1)$, there exists a $C_\beta > 0$ independent of $r > 0$ and Λ such that

$$\big\| A_\Lambda - A_\Lambda^{(N)} \,\big|\, H^1(\mathbb{R}^3) \to H^{-1}(\mathbb{R}^3) \big\| \leq C_\beta \big(r^{(1-\beta)} + r^{-1} \exp(-\pi\sqrt{N/2}) \big)\,. \tag{5.10}$$

Balancing the two terms on the right hand side, to achieve a total error ε we need $N \sim \log^2 \varepsilon$, where β enters only in the constant. Note that the estimate carries over directly to one-electron potentials in the many-electron case.

Our strategy will be to make use of the stability properties of the wavelet basis and combine error estimates for single entries using the following lemma, cf. [22, Section 8.4].

Lemma 5.2 (Schur's Lemma)**.** *Let $M = (m_{ij})_{i,j\in\mathbb{N}}$ be an infinite matrix and let $\omega_i > 0, i \in \mathbb{N}$. Suppose that*

$$\sum_j |m_{ij}|\omega_j \leq \omega_i\,, \quad \sum_i |m_{ij}|\omega_i \leq \omega_j\,,$$

then $M \colon \ell^2 \to \ell^2$ is bounded with $\|M\| \leq 1$.

Although in the following we allow arbitrary index sets Λ and hence derive an estimate for the full function space, which would also be possible without reference to the wavelet basis, the present approach has the advantage that it can also be adapted to specific families of Λ and correspondingly optimized separable approximations of the potential. Furthermore, if information on the approximation of derivatives of the potentials is available (which, however, is not the case here), it can be exploited using vanishing moment properties of the wavelets.

Proof. Due to our assumption that $\{2^{-|\lambda|}\gamma_\lambda\}$ is a Riesz basis for $H^1(\mathbb{R}^3)$,

$$\left\|A_\Lambda - A_\Lambda^{(N)} \,\middle|\, H^1(\mathbb{R}^3) \to H^{-1}(\mathbb{R}^3)\right\| \lesssim \left\|D_\Lambda \,\middle|\, \ell^2(\Lambda) \to \ell^2(\Lambda)\right\|,$$

where $D_\Lambda = (d_{\nu\mu})_{\nu,\mu\in\Lambda}$ with $d_{\nu\mu} := 2^{-|\mu|-|\nu|}|a_{\nu\mu} - a_{\nu\mu}^{(N)}|$.

Let $s_\lambda = \operatorname{supp}\gamma_\lambda$, $s_{\mu\nu} = s_\mu \cap s_\nu$ and define $\chi(\mu,\nu)$ to be one if $s_{\mu\nu} \neq \emptyset$ and zero otherwise. For $r > 0$, let $\Lambda_r := \{\lambda \in \Lambda\colon s_\lambda \cap B_r(0) \neq \emptyset\}$. Making use of nonnegativity of matrix entries, we estimate $\|D_\Lambda\| \le \|D_\Lambda^{(0)}\| + \|D_\Lambda^{(1)}\|$, where $D_\Lambda^{(0)} = (d_{\nu\mu}^{(0)})$ and $D_\Lambda^{(1)} = (d_{\nu\mu}^{(1)})$ with

$$d_{\nu\mu}^{(0)} := \begin{cases} d_{\nu\mu}\,, & (\nu,\mu) \in (\Lambda_r)^2\,, \\ 0\,, & \text{otherwise,} \end{cases} \qquad d_{\nu\mu}^{(1)} := r^{-1}\delta(N)2^{-|\mu|-|\nu|}\int |\gamma_\mu\gamma_\nu|\,dx\,.$$

At several points, we will make use of the simple estimate[3]

$$\|\gamma_\mu\gamma_\nu\|_{L^p} \lesssim 2^{-(3/p-3/2)(|\mu|\vee|\nu|)}\,2^{(3/2)(|\mu|\wedge|\nu|)}\,.$$

For the elements of $D_\Lambda^{(1)}$, with $p = 1$ the above gives the estimate

$$d_{\nu\mu}^{(1)} \lesssim r^{-1}\delta(N)\,2^{-|\mu|-|\nu|}\,2^{-3||\mu|-|\nu||/2}\chi(\mu,\nu)\,.$$

In this case, we can proceed using Lemma 5.2 similarly to [7, Section 41] with $\omega_\mu = 2^{-5|\mu|/2}$ to obtain

$$\begin{aligned}
&\omega_\nu^{-1}\sum_{\mu\in\Lambda}\omega_\mu 2^{-|\mu|-|\nu|}\,2^{-3||\mu|-|\nu||/2}\,\chi(\mu,\nu) \\
&\quad= \omega_\nu^{-1}\sum_{j\ge0}2^{-j-|\nu|}\,2^{-3|j-|\nu||/2}\sum_{|\mu|=j}\omega_\mu\,\chi(\mu,\nu) \\
&\quad\lesssim 2^{5|\nu|/2}\sum_{j\le|\nu|}2^{-5j/2}\,2^{-3/2(|\nu|-j)}\,2^{-|\nu|-j} + 2^{5|\nu|/2}\sum_{j>|\nu|}2^{-5j/2}\,2^{-3/2(j-|\nu|)}\,2^{-|\nu|-j}\,2^{3(j-|\nu|)} \\
&\quad= 2^{5|\nu|/2}\sum_{j\ge0}2^{-5j/2}\,2^{3(j-|\nu|)/2}\,2^{-|\nu|-j} = \sum_{j\ge0}2^{-2j} < \infty\,,
\end{aligned}$$

and hence $\|D_\Lambda^{(1)}\| \lesssim r^{-1}\delta(N)$.

It remains to estimate $\|D_\Lambda^{(0)}\|$. We choose $j_0 \in \mathbb{N}$ such that $2^{-(j_0+1)} < r \le 2^{-j_0}$ and set $\tilde{B}_{j_0} := B_{2^{-j_0}}(0)$; we obtain an upper estimate for the norm by considering the matrix $\tilde{D}_\Lambda^{(0)}$ with entries $d_{\mu\nu}$ for $(\mu,\nu) \in (\Lambda_{2^{-j_0}})^2$, and zero otherwise. On $\tilde{B}_{j_0}$, the pointwise approximation error in the potential is estimated by $1/|x|$.

For estimating the nonzero entries of $D_\Lambda^{(0)}$, we first consider the case $|\nu| \le j_0$. If also $|\mu| \le j_0$,

$$\left|\int_{\tilde{B}_{j_0}}\frac{1}{|x|}\gamma_\mu\gamma_\nu\,dx\right| \le \left(\int_{\tilde{B}_{j_0}}\frac{1}{|x|}\,dx\right)\sup|\gamma_\mu\gamma_\nu| \lesssim 2^{-2j_0}\,2^{3(|\mu|+|\nu|)|/2}\,,$$

and consequently

$$\tilde{d}_{\nu\mu}^{(0)} \lesssim 2^{-|\nu|-|\mu|}\,2^{-2j_0}\,2^{3(|\mu|+|\nu|)|/2} \le 2^{-j_0}\,2^{-|\nu|/2+|\mu|/2}\,. \tag{5.11}$$

If $|\mu| > j_0$ and $0 \in s_\mu$ we obtain $\tilde{d}_{\nu\mu}^{(0)} \lesssim 2^{-3|\mu|/2+|\nu|/2}$ analogously to (5.11). If $|\mu| > j_0$ and $0 \notin s_\mu$, there is a $k \in \mathbb{N}$ such that $k2^{-|\mu|} \le \operatorname{dist}(s_\mu,0) \le (k+1)2^{-|\mu|}$. The number of such wavelet indices μ for given k is of order $\mathcal{O}(k^2)$, and each individual entry can be estimated by

$$\tilde{d}_{\nu\mu}^{(0)} \le 2^{-|\mu|-|\nu|}\int_{\tilde{B}_{j_0}\cap s_\mu}\frac{1}{|x|}|\gamma_\mu\gamma_\nu|\,dx \lesssim 2^{-|\mu|-|\nu|}\frac{2^{|\mu|}}{k}2^{-3(|\mu|-|\nu|)/2} = \frac{2^{-3|\mu|/2+|\nu|/2}}{k}\,.$$

[3]Here $a \vee b = \max(a,b)$, $a \wedge b = \min(a,b)$.

In the case $|\nu| > j_0$, if $|\mu| \le |\nu|$ we consider the two cases $0 \in s_\nu$, where we get $\tilde{d}^{(0)}_{\nu\mu} \lesssim 2^{-3|\nu|/2+|\mu|/2}$, and $0 \notin s_\nu$, which also gives $\tilde{d}^{(0)}_{\nu\mu} \lesssim 2^{|\nu|}\, 2^{-3(|\nu|-|\mu|)/2}\, 2^{-|\nu|-|\mu|} = 2^{-3|\nu|/2+|\mu|/2}$. If $|\mu| > |\nu|$ and $0 \in s_\mu$, we obtain $\tilde{d}^{(0)}_{\nu\mu} \lesssim 2^{-3|\mu|/2+|\nu|/2}$, whereas if $|\mu| > |\nu|$ and $0 \notin s_\mu$, we again classify indices μ according to $\operatorname{dist}(s_\mu, 0)$ to obtain

$$\tilde{d}^{(0)}_{\nu\mu} \lesssim 2^{-|\mu|-|\nu|}\frac{2^{|\mu|}}{k} 2^{-3(|\mu|-|\nu|)/2} = \frac{2^{-3|\mu|/2+|\nu|/2}}{k}$$

with $\mathcal{O}(k^2)$ such indices μ for given k.

We apply Lemma 5.2 with $\omega_\mu = 2^{-\alpha|\mu|}$, $\alpha \in (1/2, 3/2]$. In the case $|\nu| \le j_0$,

$$\begin{aligned}
\omega_\nu^{-1} \sum_{\mu\in\Lambda} \omega_\mu |\tilde{d}^{(0)}_{\nu\mu}| &\lesssim 2^{\alpha|\nu|} \sum_{j\le j_0} 2^{-\alpha j}\, 2^{-j_0}\, 2^{-|\nu|/2+j/2} \\
&\quad + 2^{\alpha|\nu|} \sum_{j>j_0} 2^{-\alpha j} \left(2^{-3j/2+|\nu|/2} + 2^{-3j/2+|\nu|/2} \sum_{k=1}^{2^{j-j_0}} (k^2 k^{-1}) \right) \\
&\lesssim 2^{-j_0} 2^{(\alpha-1/2)|\nu|} \sum_{j\le j_0} 2^{-(\alpha-1/2)j} + 2^{(\alpha+1/2)|\nu|} \sum_{j>j_0} 2^{-(3/2+\alpha)j} \\
&\quad + 2^{(\alpha+1/2)|\nu|} \sum_{j>j_0} 2^{-(3/2+\alpha)j}\, 2^{2(j-j_0)}\,,
\end{aligned}$$

and making use of $|\nu| \le j_0$,

$$\begin{aligned}
&\lesssim 2^{-j_0}\, 2^{(\alpha-1/2)|\nu|} \sum_{j\ge 0} 2^{-(\alpha-1/2)j} + 2^{(\alpha+1/2)|\nu|}\, 2^{-(3/2+\alpha)j_0} \\
&\lesssim 2^{-(3/2-\alpha)j_0} + 2^{-j_0}\,,
\end{aligned}$$

where the constants appearing in the last steps depend on $\alpha > 1/2$. In the second case $|\nu| > j_0$,

$$\begin{aligned}
\omega_\nu^{-1} \sum_{\mu\in\Lambda} \omega_\mu |\tilde{d}^{(0)}_{\nu\mu}| &\lesssim 2^{\alpha|\nu|} \sum_{j\le|\nu|} 2^{-\alpha j}\, 2^{-3/2|\nu|+j/2} \\
&\quad + 2^{\alpha|\nu|} \sum_{j>|\nu|} 2^{-\alpha j} \left(2^{-3j/2+|\nu|/2} + 2^{-3j/2+|\nu|/2} \sum_{k=1}^{2^{j-|\nu|}} (k^2 k^{-1}) \right) \\
&\lesssim 2^{-(3/2-\alpha)|\nu|} \sum_{j\le|\nu|} 2^{-(\alpha-1/2)j} + 2^{(\alpha+1/2)|\nu|} \sum_{j>|\nu|} 2^{-(3/2+\alpha)j} \\
&\quad + 2^{(\alpha+1/2)|\nu|} \sum_{j>|\nu|} 2^{-(3/2+\alpha)j}\, 2^{2(j-|\nu|)} \\
&= 2^{-(3/2-\alpha)|\nu|} \sum_{j\ge 0} 2^{-(\alpha-1/2)j} + 2^{(\alpha+1/2)|\nu|} \sum_{j>|\nu|} 2^{-(3/2+\alpha)j} \\
&\lesssim 2^{-(3/2-\alpha)|\nu|} + 2^{-|\nu|} < 2^{-(3/2-\alpha)j_0} + 2^{-j_0}\,.
\end{aligned}$$

Putting this together and replacing $3/2-\alpha$ with $\alpha > 1/2$ by $1-\beta$ with $\beta > 0$, we obtain

$$\|D_\Lambda\| \le \tilde{C}_\beta \big(2^{-(1-\beta)j_0} + r^{-1} \exp(-\pi\sqrt{N/2}) \big)\,,$$

where $\tilde{C}_\beta$ also depends on the wavelet basis and tends to infinity as $\beta \to 0$. Since $2^{-j_0} \le 2r$ by choice of j_0, this yields (5.10). □

5.3.2 Approximation of Two-Electron Operators

For the two-electron operator $|x-y|^{-1}(x-y)\cdot(\nabla_x-\nabla_y)$, a slightly different type of exponential sum approximation seems more appropriate. The result in [4, Lemma 2, Remark 3] yields the existence of an exponential sum approximation

$$\sqrt{t}\left|\frac{1}{\sqrt{t}}-\sum_{k=1}^{N}\omega_k e^{-\alpha_k t}\right|\le c_0\frac{R}{r}\exp(-\pi\sqrt{N})=:\delta(N,r,R)\,,\quad r^2\le t\le R^2\,,\tag{5.12}$$

which for $r\le|x-y|\le R$ yields a separable approximation

$$\left|\frac{x-y}{|x-y|}-\sum_{k=1}^{N}(x-y)\omega_k e^{-\alpha_k|x-y|^2}\right|_{\ell^2}\le|x-y|\left|\frac{1}{|x-y|}-\sum_{k=1}^{N}\omega_k e^{-\alpha_k|x-y|^2}\right|\le\delta(N,r,R)\,.\tag{5.13}$$

Theorem 5.3. *Let $\Lambda\subseteq(\bar\Lambda^{(3)})^2$ such that*

$$\sup\big\{|x-y|\colon x,y\in\operatorname{supp}\Psi_\lambda\,,\ \lambda\in\Lambda\big\}\le R\,,\quad J:=\max_{\lambda\in\Lambda}|\lambda|<\infty$$

with $R>0$ sufficiently large, and let $B_\Lambda, B_\Lambda^{(N)}\colon H^1(\mathbb{R}^6)\to H^{-1}(\mathbb{R}^6)$ be the operators defined by the matrices $(b_{\nu\mu})_{\nu,\mu\in\Lambda}$, $(b_{\nu\mu}^{(N)})_{\nu,\mu\in\Lambda}$ with

$$\begin{aligned}b_{\nu\mu}&=\int_{\mathbb{R}^6}\frac{x-y}{|x-y|}\cdot\big(\nabla_x-\nabla_y\big)\big(\gamma_{\mu_1}(x)\gamma_{\mu_2}(y)\big)\,\gamma_{\nu_1}(x)\gamma_{\nu_2}(y)\,d(x,y)\,,\\ b_{\nu\mu}^{(N)}&=\sum_{k=1}^{N}\omega_k\int_{\mathbb{R}^6}(x-y)e^{-\alpha_k|x-y|^2}\big(\nabla_x-\nabla_y\big)\big(\gamma_{\mu_1}(x)\gamma_{\mu_2}(y)\big)\,\gamma_{\nu_1}(x)\gamma_{\nu_2}(y)\,d(x,y)\,,\end{aligned}$$

where $\{\omega_k\}$, $\{\alpha_k\}$ satisfy (5.13) with $0<r<R$, then there exists a $C>0$ independent of r,R and Λ such that

$$\big\|B_\Lambda-B_\Lambda^{(N)}\,\big|\,H^1(\mathbb{R}^6)\to H^{-1}(\mathbb{R}^6)\big\|\le C\big(J2^{2J}r^3+J^2r^{-1}R\exp(-\pi\sqrt{N})\big)\,.\tag{5.14}$$

Note that despite the dependence on J, if $r\lesssim 2^{-J}$ this still resembles the estimate for the one-electron potential.

Proof. We estimate the norm of $E_\Lambda\colon\ell^2(\Lambda)\to\ell^2(\Lambda)$ with entries

$$e_{\nu\mu}:=(2^{2|\mu_1|}+2^{2|\mu_2|})^{-1/2}(2^{2|\nu_1|}+2^{2|\nu_2|})^{-1/2}|b_{\mu\nu}-b_{\mu\nu}^{(N)}|\,.$$

Similarly to above, we estimate $\|E_\Lambda\|\le\|E_\Lambda^{(0)}\|+\|E_\Lambda^{(1)}\|$ where, again using the notation $s_{\mu\nu}=\operatorname{supp}\gamma_\mu\cap\operatorname{supp}\gamma_\nu$,

$$\begin{aligned}e_{\nu\mu}^{(0)}&:=\begin{cases}e_{\mu\nu}\,, & \operatorname{dist}(s_{\mu_1\nu_1},s_{\mu_2\nu_2})\le r\,,\\ 0\,, & \text{otherwise,}\end{cases}\\ e_{\nu\mu}^{(1)}&:=\delta\,(2^{2|\mu_1|}+2^{2|\mu_2|})^{-1/2}(2^{2|\nu_1|}+2^{2|\nu_2|})^{-1/2}\int|(\nabla_x-\nabla_y)\gamma_{\mu_1}\gamma_{\mu_2}|\,|\gamma_{\nu_1}\gamma_{\nu_2}|\,d(x,y)\,.\end{aligned}$$

Due to nonsymmetry, here the two estimates required in Lemma 5.2 do not coincide.

The nonzero entries of $E_\Lambda^{(1)}$ are estimated by

$$\begin{aligned}e_{\nu\mu}^{(1)}&\lesssim\delta(2^{2|\mu_1|}+2^{2|\mu_2|})^{-1/2}(2^{2|\nu_1|}+2^{2|\nu_2|})^{-1/2}\,(2^{|\mu_1|}+2^{|\mu_2|})\,2^{-3||\mu_1|-|\nu_1||/2}\,2^{-3||\mu_2|-|\nu_2||/2}\\ &\lesssim\delta 2^{-(|\nu_1|\vee|\nu_2|)}\,2^{-3||\mu_1|-|\nu_1||/2}\,2^{-3||\mu_2|-|\nu_2||/2}\,,\\ e_{\mu\nu}^{(1)}&\lesssim\delta 2^{-(|\mu_1|\vee|\mu_2|)}\,2^{-3||\mu_1|-|\nu_1||/2}\,2^{-3||\mu_2|-|\nu_2||/2}\,.\end{aligned}$$

In the first case of $e^{(1)}_{\nu\mu}$, using weights $\omega_{|\mu_1|,|\mu_2|} = 2^{-3(|\mu_1|+|\mu_2|)/2}$,

$$\omega^{-1}_{|\nu_1|,|\nu_2|}\sum_{\mu}\omega_{|\mu_1|,|\mu_2|}|e^{(1)}_{\nu\mu}| \lesssim \omega^{-1}_{|\nu_1|,|\nu_2|}\,2^{-3(|\nu_1|+|\nu_2|)/2}\,2^{-(|\nu_1|\vee|\nu_2|)}\delta\sum_{j_1,j_2}\omega_{j_1,j_2}\,2^{3(j_1+j_2)/2}$$
$$= 2^{-(|\nu_1|\vee|\nu_2|)}\delta\sum_{j_1,j_2}1 \;\sim\; 2^{-(|\nu_1|\vee|\nu_2|)}\,\delta J^2\,,$$

whereas in the second case of the transposed $e^{(1)}_{\mu\nu}$,

$$\omega^{-1}_{|\nu_1|,|\nu_2|}\sum_{\mu}\omega_{|\mu_1|,|\mu_2|}|e^{(1)}_{\mu\nu}| \lesssim \omega^{-1}_{|\nu_1|,|\nu_2|}\,2^{-3(|\nu_1|+|\nu_2|)/2}\,\delta\sum_{j_1,j_2}\omega_{j_1,j_2}\,2^{-(j_1\vee j_2)}\,2^{3(j_1+j_2)/2}$$
$$= \delta\sum_{j_1,j_2}2^{-(j_1\vee j_2)} = \delta\sum_{j}(2j+1)2^{-j}\,,$$

where the sum is bounded independently of J. Altogether, this gives $\|E^{(1)}_\Lambda\| \le CJ^2\delta$.

For $E^{(0)}_\Lambda$, we again take j_0 such that $2^{-(j_0+1)} < r \le 2^{-j_0}$. Without loss of generality, we can assume $|\nu_1| \ge |\nu_2|$. First we observe that given such ν, for any $0 \le j_1, j_2 \le J$, with $\Lambda_{j_1,j_2} := \{\mu\in\Lambda\colon |\mu_1| = j_1, |\mu_2| = j_2\}$,

$$\sum_{\mu\in\Lambda_{j_1,j_2}}\big|\{|x-y|\le 2^{-j_0}\}\cap(s_{\mu_1\nu_1}\times s_{\mu_2\nu_2})\big| \lesssim 2^{-3|\nu_1|}\,2^{-3(j_0\vee|\nu_2|)}\,.$$

Hence, with $\omega_{|\mu_1|,|\mu_2|} := 2^{-4(|\mu_1|\vee|\mu_2|)}\,2^{3(|\mu_1|+|\mu_2|)/2}$,

$$\omega^{-1}_{|\nu_1|,|\nu_2|}\sum_{\mu}\omega_{|\mu_1|,|\mu_2|}|e^{(0)}_{\nu\mu}| \lesssim \omega^{-1}_{|\nu_1|,|\nu_2|}\sum_{j_1,j_2}\omega_{j_1,j_2}\,2^{-|\nu_1|}\,2^{-3|\nu_1|}\,2^{-3(j_0\vee|\nu_2|)}\,2^{3(j_1+|\nu_1|+j_2+|\nu_2|)/2}$$
$$= 2^{-3(j_0\vee|\nu_2|)}\sum_{j_1,j_2}2^{-4(j_1\vee j_2)}\,2^{3(j_1+j_2)}$$
$$= 2^{-3(j_0\vee|\nu_2|)}\sum_{j_1,j_2}2^{-(j_1\vee j_2)+3(j_1\wedge j_2)}$$
$$\le 2^{-3(j_0\vee|\nu_2|)}\sum_{j_1,j_2}2^{2(j_1\vee j_2)} \;\lesssim\; 2^{-3j_0}J2^{2J}\,.$$

For the transposed $e^{(0)}_{\mu\nu}$,

$$\omega^{-1}_{|\nu_1|,|\nu_2|}\sum_{\mu}\omega_{|\mu_1|,|\mu_2|}|e^{(0)}_{\mu\nu}| \lesssim \omega^{-1}_{|\nu_1|,|\nu_2|}\sum_{j_1,j_2}\omega_{j_1,j_2}\,2^{-(j_1\vee j_2)}\,2^{-3|\nu_1|}\,2^{-3(j_0\vee|\nu_2|)}\,2^{3(j_1+|\nu_1|+j_2+|\nu_2|)/2}$$
$$= 2^{-3(j_0\vee|\nu_2|)}\,2^{|\nu_1|}\sum_{j_1,j_2}2^{-5(j_1\vee j_2)}\,2^{3(j_1+j_2)}$$
$$\le 2^{-3(j_0\vee|\nu_2|)}\,2^{|\nu_1|}\sum_{j_1,j_2}2^{(j_1\vee j_2)} \;\lesssim\; 2^{-3j_0}J2^{2J}\,.$$

Combining both parts, we arrive at (5.14). □

6 Conclusion

We have investigated the mixed regularity of eigenfunctions arising in a transcorrelated method in the two-electron case and discussed some implementational aspects for hyperbolic wavelet discretizations, in particular deriving estimates in operator norm for separable approximation of potentials, which shows a path towards practically feasible methods based on such concepts.

There are many directions for further work. In particular, it remains to combine the different approximation estimates to a complete analysis of a discretization scheme based on the transcorrelated formulation. Due to nonsymmetry of the modified problem and the lack of compact embedding of the involved function spaces, this is not covered by standard results. A complete analysis, further details concerning a practical realization, and numerical experiments will be the subject of a forthcoming publication.

The eventual aim is, of course, to approach problems with several electrons. In this case, the nonsymmetric formulation shows its major advantage, compared to other correlated treatments, of not requiring integrals over four or more electron coordinates.

Acknowledgements. Financial support from Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111 is gratefully acknowledged. The author is thankful to Joachim Schöberl for useful discussions; to Karen Veroy and Wolfgang Dahmen for their advice, fruitful discussions on the subject, and helpful comments concerning the presentation in this work; and to Edoardo Di Napoli and Anna-Lena Gerner for their valuable comments.

References

[1] Agmon, S. *Lectures on Exponential Decay of Solutions of Second-Order Elliptic Equations: Bounds on Eigenfunctions of N-Body Schrödinger Operators*. Mathematical Notes. Princeton University Press, 1982.

[2] Beylkin, G. On the representation of operators in bases of compactly supported wavelets. *SIAM Journal on Numerical Analysis 29*, 6 (December 1992), 1716–1740.

[3] Boys, S. F., and Handy, N. C. The determination of energies and wavefunctions with full electronic correlation. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 310*, 1500 (1969), 43–61.

[4] Braess, D. Asymptotics for the approximation of wave functions by exponential sums. *Journal of Approximation Theory 83* (1995), 93–103.

[5] Braess, D., and Hackbusch, W. On the efficient computation of high-dimensional integrals and the approximation by exponential sums. In *Multiscale, Nonlinear and Adaptive Approximation*, R. DeVore and A. Kunoth, Eds. Springer Berlin Heidelberg, 2009.

[6] Chinnamsetty, S. R., Espig, M., Khoromskij, B. N., Hackbusch, W., and Flad, H.-J. Tensor product approximation with optimal rank in quantum chemistry. *The Journal of chemical physics 127*, 8 (2007), 084110.

[7] Cohen, A. *Wavelet methods in numerical analysis*, vol. VII of *Handbook of Numerical Analysis*. Elsevier, 2000, pp. 417–711.

[8] Dahmen, W., and Micchelli, C. A. Using the refinement equation for evaluating integrals of wavelets. *SIAM Journal on Numerical Analysis 30*, 2 (April 1993), 507–537.

[9] Daubechies, I. *Ten Lectures on Wavelets*. CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, 1992.

[10] Dijkema, T. J., Schwab, C., and Stevenson, R. An adaptive wavelet method for solving high-dimensional elliptic PDEs. *Constructive Approximation 30*, 3 (2009), 423–455. Preprint.

[11] Donovan, G., Geronimo, J., and Hardin, D. Intertwining multiresolution analyses and the construction of piecewise-polynomial wavelets. *SIAM Journal on Mathematical Analysis 27*, 6 (1996), 1791–1815.

[12] Flad, H.-J., Hackbusch, W., Khoromskij, B. N., and Schneider, R. Concepts of data-sparse tensor-product approximation in many-particle modelling. Tech. Rep. 3/08, MPI MIS, Leipzig, 2008.

[13] Flad, H.-J., Hackbusch, W., Kolb, D., and Schneider, R. Wavelet approximation of correlated wave functions. I. Basics. *Journal of Chemical Physics 116*, 22 (June 2002).

[14] Flad, H.-J., Hackbusch, W., and Schneider, R. Best N-term approximation in electronic structure calculations. II. Jastrow factors. *ESAIM Mathematical Modelling and Numerical Analysis 41*, 2 (2007), 261.

[15] Fournais, S., Hoffmann-Ostenhof, M., Hoffmann-Ostenhof, T., and Østergaard Sørensen, T. Sharp regularity results for many-electron wave functions. *Communications in Mathematical Physics 255* (2005), 183–227.

[16] Gilbarg, D., and Trudinger, N. S. *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics. Springer Berlin Heidelberg, 1998.

[17] Hirschfelder, J. O. Removal of electron-electron poles from many-electron Hamiltonians. *The Journal of Chemical Physics 39*, 11 (1963), 3145–3146.

[18] Hylleraas, E. Über den Grundzustand des Heliumatoms. *Z. Phys. 48* (1929), 469.

[19] Kato, T. On the eigenfunctions of many-particle systems in quantum mechanics. *Comm. Pure Appl. Math. X* (1957), 151–177.

[20] Klopper, W. R12 methods, gaussian geminals. In *Moderns Methods and Algorithms of Quantum Chemistry* (Jülich, 2000), J. Grotendorst, Ed., pp. 181–229.

[21] Luo, H., Kolb, D., Flad, H.-J., Hackbusch, W., and Koprucki, T. Wavelet approximation of correlated wave functions. II. Hyperbolic wavelets and adaptive approximation schemes. *Journal of Chemical Physics 117*, 8 (August 2002).

[22] Meyer, Y., and Coifman, R. *Wavelets: Calderon-Zygmund and multilinear operators*. Cambridge studies in advanced mathematics. Cambridge University Press, 1997.

[23] Nooijen, M., and Bartlett, R. J. Elimination of Coulombic infinities through transformation of the Hamiltonian. *J. Chem. Phys. 109*, 19 (1998).

[24] Reed, M., and Simon, B. *Methods of Modern Mathematical Physics, IV: Analysis of Operators*. Academic Press, 1978.

[25] Schwab, C., and Todor, R. A. Sparse finite elements for stochastic elliptic problems – higher order moments. *Computing 71*, 1 (2003), 43–63.

[26] Sweldens, W. The lifting scheme: A custom-design construction of biorthogonal wavelets. *Applied and Computational Harmonic Analysis 3*, 2 (1996), 186–200.

[27] Ten-no, S. A feasible transcorrelated method for treating electronic cusps using a frozen Gaussian geminal. *Chemical Physics Letters 330* (2000), 169–174.

[28] Tew, D. P., and Klopper, W. New correlation factors for explicitly correlated electronic wave functions. *J Chem Phys 123*, 7 (2005), 074101.

[29] Yserentant, H. On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives. *Numerische Mathematik 98* (August 2004), 731–759.

[30] Yserentant, H. The mixed regularity of electronic wave functions multiplied by explicit correlation factors. DFG SPP 1324 Preprint 49, June 2010.

[31] Yserentant, H. *Regularity and Approximability of Electronic Wave Functions*, vol. 2000 of *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, 2010.

[32] Zweistra, H. J. A., Samson, C. C. M., and Klopper, W. Similarity-transformed hamiltonians by means of gaussian-damped interelectronic distances. *Collection of Czechoslovak Chemical Communications 68*, 2 (2003), 374–386.