ESAIM: M2AN 46 (2012) 1337–1362
DOI: 10.1051/m2an/2012009

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

HYPERBOLIC WAVELET DISCRETIZATION OF THE TWO-ELECTRON SCHRÖDINGER EQUATION IN AN EXPLICITLY CORRELATED FORMULATION

Markus Bachmayr[1]

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 and the improvement compared to the standard formulation are discussed. Elements of a discretization of the eigenvalue problem based on orthogonal wavelets are described, and possible choices of tensor product bases are compared especially from an algorithmic point of view. The use of separable approximations of potential terms for applying operators efficiently is studied in detail, and estimates for the error due to this further approximation are given.

Mathematics Subject Classification. 35B65, 35J10, 65T60, 81Q05.

Received June 18, 2010. Revised May 9, 2011.
Published online March 30, 2012.

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 considered here, it is an eigenvalue problem which has as its solutions the possible quantum states 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 $u\colon (\mathbb{R}^3 \times \{-\frac{1}{2}, \frac{1}{2}\})^n \to \mathbb{R}$,

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

For each electron, we have coordinates $(x, \sigma) \in \mathbb{R}^3 \times \{-\frac{1}{2}, \frac{1}{2}\}$, 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

Keywords and phrases. Schrödinger equation, mixed regularity, transcorrelated method, wavelets, separable approximation.

[1] RWTH Aachen, Institut für Geometrie und Praktische Mathematik, Templergraben 55, 52056 Aachen, Germany.
bachmayr@igpm.rwth-aachen.de

Article published by EDP Sciences

© EDP Sciences, SMAI 2012

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

(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.

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 [34]; in the form given in [20] it amounts to the assertion that eigenfunctions u of (1.2) can be written as

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

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 [18] 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 [33], 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

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

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. 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.* [47].

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. The eigenfunction regularity results of Yserentant [48,49] show that it is in principle possible to obtain a discretization with controlled error based on sparse grid- or hyperbolic wavelet-type approximation

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

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. Our objective is to advance in this direction.

In the regularity results for n-electron eigenfunctions of (1.1) collected in [49], and the recent improvement in [37], the obstacle for higher regularity – and thus for a further reduction of approximation complexity – lies in 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 [48] 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 prerequisites for computational schemes. This concerns in particular the efficient approximate application of operators, and the related question of suitable tensor product wavelet bases.

After completion of the regularity results in Section 3, we have become aware of the recent work by Yserentant [50] that contains these results as a special case. Since it allows some substantial simplifications as compared to the general proof for n electrons in [50], and highlights the role of explicit correlation more clearly for our further discussion, we include our short proof for the two-electron case.

2. Transcorrelated formulation

In this section, we consider 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 $u, v \in \mathrm{H}^1(\mathbb{R}^{3n})$ as

$$a(u,v) := \frac{1}{2}\int \mathrm{D}u \cdot \mathrm{D}v \,\mathrm{d}x + \int (V_{\mathrm{ne}} + V_{\mathrm{ee}})\, u\, v \,\mathrm{d}x$$

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

$$a(u,v) = \lambda \langle u, v\rangle \quad \text{for all } v \in \mathrm{H}^1(\mathbb{R}^{3n}), \tag{2.1}$$

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

$$F(x) := \frac{1}{2}\sum_{i<j} |x_i - x_j|. \tag{2.2}$$

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.

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

The approach we will follow here corresponds to taking $u = \exp(F)\varphi$, $v = \exp(-F)\tau$ 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.

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 especially for the two-electron case is not a major issue in our setting. Therefore, and to avoid more complicated expressions, we use the unbounded correlation factor $\exp(F)$ in what follows.

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

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

Written out in full, we have

$$\begin{aligned}\tilde{a}(\varphi,\tau) = {}& \frac{1}{2}\int \mathrm{D}\varphi\cdot\mathrm{D}\tau\,\mathrm{d}x - \sum_{i,\nu}\int \frac{Z_\nu}{|x_i-a_\nu|}\varphi\,\tau\,\mathrm{d}x \\ & - \frac{1}{2}\sum_i \int \sum_{k\neq i}\frac{x_i-x_k}{|x_i-x_k|}\cdot \mathrm{D}_{x_i}\varphi\,\tau\,\mathrm{d}x - \frac{1}{8}\sum_i\int\sum_{k,l\neq i}\frac{x_i-x_k}{|x_i-x_k|}\cdot\frac{x_i-x_l}{|x_i-x_l|}\varphi\,\tau\,\mathrm{d}x. \end{aligned} \tag{2.3}$$

The strong form of the modified problem (2.3) was also used to obtain the regularity results in [20] already mentioned above. In the quantum chemistry literature, the formulation seems to appear first in [32]; it was used in computational schemes for Gaussian-type orbitals for instance in [6, 41, 46, 54]. Quite promising numerical results using Gaussian basis sets for Helium for the particular Hamiltonian corresponding to (2.3) are given in [36].

We begin by establishing the connection between a and $\tilde{a}$, see also [50].

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

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

Proof. From Lipschitz continuity of $\mathrm{e}^{-|\cdot|}$ and the chain and product rules for weak differentiation (*cf.* [23]) it follows that $\mathrm{e}^{-F}u \in \mathrm{H}^1(\mathbb{R}^{3n})$. For any $\tau \in \mathcal{D}(\mathbb{R}^{3n})$,

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

Now on the one hand,

$$\frac{1}{2}\int \mathrm{e}^{-F}\mathrm{D}u\cdot\mathrm{D}\tau\,\mathrm{d}x - \frac{1}{2}\int \mathrm{e}^{-F}\mathrm{D}F\cdot\mathrm{D}u\,\tau\,\mathrm{d}x = \frac{1}{2}\int \mathrm{D}u\cdot\mathrm{D}(\mathrm{e}^{-F}\tau)\,\mathrm{d}x,$$

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

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

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

Putting this together, using that $\mathrm{e}^{-F}\tau \in \mathrm{H}^1(\mathbb{R}^{3n})$ and that u solves (2.1),

$$\tilde{a}(\mathrm{e}^{-F}u,\tau) = \frac{1}{2}\int \mathrm{D}u\cdot\mathrm{D}(\mathrm{e}^{-F}\tau)\,\mathrm{d}x + \int \mathrm{e}^{-F}(V_{\mathrm{ne}}+V_{\mathrm{ee}})u\,\tau\,\mathrm{d}x = a(u,\mathrm{e}^{-F}\tau) = \lambda\langle u,\mathrm{e}^{-F}\tau\rangle = \lambda\langle \mathrm{e}^{-F}u,\tau\rangle,$$

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

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

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

The decay behavior of eigenfunctions of Schrödinger operators has been studied quite extensively, see the monograph [2]. We quote a result given in [49] that is most appropriate for our purposes.

Theorem 2.2. *If $u\in\mathrm{H}^1(\mathbb{R}^{3n})$ is an eigenfunction belonging to an eigenvalue λ in the discrete spectrum of (1.1), then for any $\delta>0$ with $\delta<\sqrt{2d_\lambda}$,*

$$\int \mathrm{e}^{2\delta|x|}\left(|u|^2+|\mathrm{D}u|^2\right)\,\mathrm{d}x \;<\infty, \tag{2.6}$$

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 $w^*=\mathrm{e}^F u\in\mathrm{H}^1(\mathbb{R}^{3n})$, 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 $\tilde{a}(\tau,\mathrm{e}^F u)$ for $\tau\in\mathcal{D}(\mathbb{R}^{3n})$, using the integration by parts

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

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

Remark 2.4. To give a specific example of what the exponential decay property (2.6) means for w^* from Proposition 2.3 with F as in (2.2), we consider Helium, where $\Sigma=-2$ (*cf.* [42], XIII.3.A, note the different scaling convention for the Laplacian term in the Hamiltonian) and for the ground state we have $\lambda_0<-2.9037$. Since $\sqrt{2d_{\lambda_0}}|x|-\frac{1}{2}|x_1-x_2|\geq\left(\sqrt{2d_{\lambda_0}}-\frac{1}{\sqrt{2}}\right)|x|$, estimate (2.6) is satisfied for w^* with $\delta<\sqrt{2d_{\lambda_0}}-\frac{1}{\sqrt{2}}\approx 0.637$.

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 $f\in\mathrm{L}^2(\mathbb{R}^{3n})$ for which

$$\left\|f\,|\,\mathcal{H}^{s,k}(\mathbb{R}^3;n)\right\|^2 := \int\left(1+\prod_i|\xi_i|^{2s}\right)\left(1+\sum_i|\xi_i|^2\right)^k|\hat{f}|^2\,\mathrm{d}(\xi_1,\ldots,\xi_n)<\infty, \tag{3.1}$$

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

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 $\mathrm{H}^k(\mathbb{R}^{3n})$.

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

$$-\frac{1}{2}\Delta u + V_{\mathrm{ne}} u + \frac{1}{|x-y|} u = \lambda u, \tag{3.2}$$

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

$$Tw := -\frac{1}{2}\Delta w + V_{\mathrm{ne}}\, w - \frac{1}{2}\frac{x-y}{|x-y|}\cdot(\mathrm{D}_x - \mathrm{D}_y)w - \frac{1}{4}w = \lambda w \tag{3.3}$$

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

This regularity statement for w cannot be expected to be sharp. For u, it is shown in [48] by similar arguments as adapted here that $u \in \mathcal{H}^{1/2,1}(\mathbb{R}^3; 2)$; in [37], this is improved to the sharp result $u \in \mathcal{H}^{s,1}(\mathbb{R}^3; 2)$ for $s < \frac{3}{4}$.

On the spaces $\mathcal{H}^{1,k}(\mathbb{R}^3; 2)$, $k = 0, 1$, which can be identified with closures of $\mathcal{D}(\mathbb{R}^6)$ under the norms defined in (3.1), we make use of the equivalent norms and seminorms

$$\|u\|_{1,k} = \left(\|u\|_k^2 + \sum_{i,j=1}^3 \|\mathrm{D}^2_{x_i,y_j} u\|_k^2\right)^{1/2}, \quad |u|_{1,0} = \left(\sum_{i,j=1}^3 \|\mathrm{D}^2_{x_i,y_j} u\|_0^2\right)^{1/2}, \quad |u|_{1,1} = \left(\sum_{i,j=1}^3 |\mathrm{D}^2_{x_i,y_j} u|_1^2\right)^{1/2}.$$

We shall also use the space $\mathcal{H}^{1,-1}(\mathbb{R}^3; 2)$ defined by duality.

The proof of Theorem 3.1 requires a few preparatory results. We follow the basic strategy of [48]: First, we show boundedness and ellipticity in $\mathcal{H}^{1,1}$ of the augmented bilinear form

$$b(u, v) := \langle (T + \mu\,\mathrm{I})u, v\rangle + \sum_{i,j=1}^3 \langle \mathrm{D}^2_{x_i,y_j}(T + \mu\,\mathrm{I})u, \mathrm{D}^2_{x_i,y_j} v\rangle \tag{3.4}$$

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 (3.3).

For ellipticity and coercivity, we need appropriate estimates for the potential terms in b. For the case of the single-particle potential V_{ne}, it was shown in [48] that

$$\left|\int V_{\mathrm{ne}} u\, v \,\mathrm{d}(x,y)\right| \lesssim \|u\|_0 |v|_1, \quad \sum_{i,j=1}^3 \left|\int \mathrm{D}^2_{x_i,y_j}(V_{\mathrm{ne}} u)\, \mathrm{D}^2_{x_i,y_j} v \,\mathrm{d}(x,y)\right| \lesssim |u|_{1,0}|v|_{1,1}, \quad u, v \in \mathcal{D}(\mathbb{R}^6). \tag{3.5}$$

A new estimate is required for the modified electron-electron operator

$$\tilde{V}_{\mathrm{ee}}(x,y) := -\frac{1}{2}\frac{x-y}{|x-y|}\cdot(\mathrm{D}_x - \mathrm{D}_y) - \frac{1}{4}. \tag{3.6}$$

It is at this point that the explicit correlation comes into play: since a better estimate than for the two-electron Coulomb potential is possible, eventually one obtains higher regularity.

We make use of the following Hardy-type inequality, *cf.* [49], Section 5.4,

$$\int \frac{v^2}{|x-y|^2}\,\mathrm{d}(x,y) \le 4 \min\left\{\int |\mathrm{D}_x v|^2\,\mathrm{d}(x,y), \int |\mathrm{D}_y v|^2\,\mathrm{d}(x,y)\right\}, \quad v \in \mathcal{D}(\mathbb{R}^6). \tag{3.7}$$

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

Lemma 3.2. *Let $u, v \in \mathcal{D}(\mathbb{R}^6)$, then there exist $C, \tilde{C} > 0$ such that*

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

$$\sum_{i,j=1}^{3} \left| \int \mathrm{D}^2_{x_i,y_j} \left(\tilde{V}_{\mathrm{ee}} u\right) \mathrm{D}^2_{x_i,y_j} v \,\mathrm{d}(x,y) \right| \leq \tilde{C}\left(|u|_{1,0}|v|_{1,1} + |u|_{1,1}|v|_{1,0}\right) + \frac{1}{4}|u|_{1,0}|v|_{1,0}.$$

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

$$\sum_{i,j} \left| \int \mathrm{D}^2_{x_i,y_j} \left(\frac{x-y}{|x-y|} \cdot \mathrm{D}_x u \right) \mathrm{D}^2_{x_i,y_j} v \,\mathrm{d}(x,y) \right| \lesssim \left(\sum_{i,j} \left\| \mathrm{D}^2_{x_i,y_j} u \right\|_0^2 \right)^{\frac{1}{2}} \left(\sum_{i,j} \left| \mathrm{D}^2_{x_i,y_j} v \right|_1^2 \right)^{\frac{1}{2}} + \left(\sum_{i,j} \left| \mathrm{D}^2_{x_i,y_j} u \right|_1^2 \right)^{\frac{1}{2}} \left(\sum_{i,j} \left\| \mathrm{D}^2_{x_i,y_j} v \right\|_0^2 \right)^{\frac{1}{2}}. \quad (3.8)$$

The term with $\mathrm{D}_y u$ on the left hand side can then be treated analogously. Using the product rule on the left hand side of (3.8) gives

$$\mathrm{D}^2_{x_i,y_j} \left(\frac{x-y}{|x-y|} \cdot \mathrm{D}_x u \right) = \sum_k \left[\left(\mathrm{D}^2_{x_i,y_j} \frac{x_k - y_k}{|x-y|} \right) \mathrm{D}_{x_k} u + \left(\mathrm{D}_{y_j} \frac{x_k - y_k}{|x-y|} \right) \mathrm{D}^2_{x_k,x_i} u + \left(\mathrm{D}_{x_i} \frac{x_k - y_k}{|x-y|} \right) \mathrm{D}^2_{x_k,y_j} u + \frac{x_k - y_k}{|x-y|} \left(\mathrm{D}^3_{x_k,x_i,y_j} u \right) \right]. \quad (3.9)$$

The integrals for the rightmost term in (3.8) with third derivatives of u can be estimated directly by $|u|_{1,1}|v|_{1,0}$, since the coefficient is uniformly bounded. For the first term in (3.9), which contains second derivatives of the coefficient, we obtain

$$\begin{aligned} \sum_{i,j} \left| \int \left(\mathrm{D}^2_{x_i,y_j} \frac{x-y}{|x-y|} \right) \cdot \mathrm{D}_x u\, \mathrm{D}^2_{x_i,y_j} v \,\mathrm{d}(x,y) \right| &\lesssim \left(\int \frac{|\mathrm{D}_x u|^2}{|x-y|^2} \,\mathrm{d}(x,y) \right)^{\frac{1}{2}} \left(\sum_{i,j} \int \frac{\left| \mathrm{D}^2_{x_i,y_j} v \right|^2}{|x-y|^2} \,\mathrm{d}(x,y) \right)^{\frac{1}{2}} \\ &\lesssim \left(\sum_{j,k} \int \left| \mathrm{D}^2_{y_j,x_k} u \right|^2 \mathrm{d}(x,y) \right)^{\frac{1}{2}} \left(\sum_{k,i,j} \int \left| \mathrm{D}^3_{x_k,x_i,y_j} v \right|^2 \mathrm{d}(x,y) \right)^{\frac{1}{2}} \\ &= |u|_{1,0}|v|_{1,1}, \end{aligned}$$

where we have used (3.7). The middle terms in (3.9) can be estimated, again using (3.7),

$$\sum_{i,j} \left| \int \left(\mathrm{D}_{y_j} \frac{x-y}{|x-y|} \right) \cdot \left(\mathrm{D}_x\, \mathrm{D}_{x_i} u \right) \left(\mathrm{D}^2_{x_i,y_j} v \right) \mathrm{d}(x,y) \right| \lesssim \left(\sum_{i,k} \int \frac{1}{|x-y|^2} \left| \mathrm{D}^2_{x_i,x_k} u \right|^2 \mathrm{d}(x,y) \right)^{\frac{1}{2}} \left(\sum_{i,j} \int \left| \mathrm{D}^2_{x_i,y_j} v \right|^2 \mathrm{d}(x,y) \right)^{\frac{1}{2}} = |u|_{1,1}|v|_{1,0},$$

where the roles of x_i and y_j can be interchanged. Altogether, this shows (3.8). □

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

Proof of Theorem 3.1. Boundedness of b as in (3.4) on $\mathcal{H}^{1,1}(\mathbb{R}^3;2)$ follows with (3.5) and Lemma 3.2. Furthermore, with constants $c,\tilde{c}>0$, for $v\in\mathcal{H}^{1,1}(\mathbb{R}^3;2)$ we have

$$b(v,v)\geq\frac{1}{2}|v|_1^2-c\|v\|_0|v|_1+\left(\mu-\frac{1}{4}\right)\|v\|_0^2+\frac{1}{2}|v|_{1,1}^2-\tilde{c}|v|_{1,0}|v|_{1,1}+\left(\mu-\frac{1}{4}\right)|v|_{1,0}^2,$$

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

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

$$b(\psi,v)=(\lambda+\mu)\left[\langle w,v\rangle+\sum_{i,j}\langle\mathrm{D}^2_{x_i,y_j}w,\mathrm{D}^2_{x_i,y_j}v\rangle\right],\quad v\in\mathcal{H}^{1,1}(\mathbb{R}^3;2)$$

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

$$\langle w,v\rangle+\sum_{i,j}\langle\mathrm{D}^2_{x_i,y_j}w,\mathrm{D}^2_{x_i,y_j}v\rangle=\langle w,v+\Delta_x\Delta_y v\rangle.$$

It can be verified as in [48] 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 $v\in\mathcal{S}(\mathbb{R}^6)$, we have $\langle(T+\mu\,\mathrm{I})\psi,v+\Delta_x\Delta_y v\rangle=(\lambda+\mu)\langle w,v+\Delta_x\Delta_y v\rangle$. 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 $\varphi\in\mathrm{H}^1(\mathbb{R}^6)$. By ellipticity, this implies $\psi=w$. □

Remark 3.3. One can slightly modify the proof along the lines of [49], Chapter 5.1 and [50] 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 $\exp(\gamma(|x|+|y|))\,w\in\mathcal{H}^{1,1}(\mathbb{R}^3;2)$ for any γ, $0<\gamma<\bar{\gamma}$.

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 arising three-electron integrals, see [50] for the general case.

Remark 3.4. 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), but following the same reasoning as in [37], one obtains that the solution of the adjoint problem is in $\mathcal{H}^{s,1}(\mathbb{R}^3;2)$ for $s<\frac{3}{4}$, *i.e.*, has the same regularity as 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 (3.3), as well as appropriate choices of wavelet bases for numerical realizations.

The following will be based on an orthonormal wavelet basis for $\mathrm{L}^2(\mathbb{R})$ 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 from scaling functions at a suitable level $j_0\in\mathbb{Z}$,

$$\{\psi_\lambda\}_{\lambda\in\nabla},\quad \nabla:=\{(j_0,k,0)\colon k\in\mathbb{Z}\}\cup\{(j,k,1)\colon j\geq j_0,k\in\mathbb{Z}\}.$$

Further details on the choice of wavelets and the reasons for working with an orthonormal basis are given in Section 4.2. We further introduce the abbreviations $|\lambda|:=j$, $\mathrm{k}(\lambda):=k$, $\mathrm{s}(\lambda):=s$ for $\lambda=(j,k,s)\in\nabla$. We assume in addition that $\varphi,\psi\in\mathrm{H}^{s^*}(\mathbb{R})$ for some $s^*>2$ and ψ has $p\geq 3$ vanishing moments, which implies that

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

$\{2^{-s|\lambda|}\psi_\lambda\}_{\lambda\in\nabla}$ is a Riesz basis for $\mathrm{H}^s(\mathbb{R})$, $0 < s \le 2$, *cf.* [11], Theorem 3.7.7. At some points of what follows, these regularity assumptions could be weakened at the expense of further technical difficulties; in Section 5.2, however, we shall use that $\varphi, \psi \in \mathrm{H}^2(\mathbb{R})$. For dimensions $d \ge 1$, we define the index sets

$$\nabla^{(d)} := \{(j_0, k, 0)\} \cup \big\{(j, k, s) \colon j \ge j_0, s \in \{0,1\}^d \setminus \{0\}\big\} \subset \mathbb{Z} \times \mathbb{Z}^d \times \{0,1\}^d$$

and corresponding tensor product basis functions with isotropic support,

$$\Psi_\lambda := \bigotimes_{i=1}^{d} \psi_{(|\lambda|, \mathrm{k}_i(\lambda), \mathrm{s}_i(\lambda))}, \quad \lambda \in \nabla^{(d)}.$$

Note that $\nabla^{(1)} = \nabla$ and $\Psi_\lambda = \psi_\lambda$ for $\lambda \in \nabla^{(1)}$. For any $d \in \mathbb{N}$, $\{\Psi_\lambda\}_{\lambda\in\nabla^{(d)}}$ is an orthonormal basis of $\mathrm{L}^2(\mathbb{R}^d)$ and $\{2^{-s|\lambda|}\Psi_\lambda\}_{\lambda\in\nabla^{(d)}}$ is a Riesz basis of $\mathrm{H}^s(\mathbb{R}^d)$, $0 < s \le 2$. Furthermore, we define $\chi_d(\mu,\nu)$ to be 1 if $\operatorname{supp}\Psi_\mu \cap \operatorname{supp}\Psi_\nu \neq \emptyset$, and 0 otherwise.

4.1. Hyperbolic wavelet bases

For $d, D \in \mathbb{N}$, tensor product wavelets on $\mathbb{R}^{dD}$ with anisotropic support can be defined by

$$\Psi_\lambda := \bigotimes_{i=1}^{D} \Psi_{\lambda_i}, \quad \lambda = (\lambda_1, \ldots, \lambda_D) \in (\nabla^{(d)})^D.$$

There are two different options for the construction of basis functions for the n-electron case that are of interest in our context: choosing $d = 3$ and $D = n$, *i.e.*, tensor products of single-electron basis functions, which directly corresponds to the regularity estimates of the previous section; or choosing $d = 1$ and $D = 3n$, *i.e.*, $3n$-fold tensor products of univariate wavelet bases. The first choice will also be referred to as a partially anisotropic, the second choice as a fully anisotropic tensor product basis.

Under our above assumptions, for $\mathrm{H}^s(\mathbb{R}^{3n})$ we have the norm equivalences

$$\|u\|_s^2 \sim \sum_{\lambda\in(\nabla^{(3)})^n} \sum_{i=1}^{n} 2^{2s|\lambda_i|} |\langle u, \Psi_\lambda\rangle|^2 \sim \sum_{\lambda\in\nabla^{3n}} \sum_{i=1}^{3n} 2^{2s|\lambda_i|} |\langle u, \Psi_\lambda\rangle|^2, \quad 0 < s \le 2; \tag{4.1}$$

for the spaces $\mathcal{H}^{s,1}(\mathbb{R}^3; n)$, $0 < s \le 1$, similarly

$$\|u\|_{s,1}^2 \sim \sum_{\lambda\in(\nabla^{(3)})^n} 2^{2s\sum_{i=1}^n |\lambda_i|} \sum_{i=1}^{n} 2^{2|\lambda_i|} |\langle u, \Psi_\lambda\rangle|^2 \sim \sum_{\lambda\in\nabla^{3n}} \left(\prod_{i=0}^{n-1} \sum_{j=1}^{3} 2^{2s|\lambda_{3i+j}|}\right) \sum_{i=1}^{3n} 2^{2|\lambda_i|} |\langle u, \Psi_\lambda\rangle|^2, \tag{4.2}$$

see also [14,26], and in each case the correspondingly scaled basis functions provide Riesz bases of the respective function spaces. In what follows, for $u \in \mathrm{L}^2(\mathbb{R}^{3n})$ and Λ any set of wavelet indices, we shall denote by $u_\Lambda := \sum_{\lambda\in\Lambda} \langle u, \Psi_\lambda\rangle \Psi_\lambda$ the L^2-orthogonal projection onto $\operatorname{span}\{\Psi_\lambda\}_{\lambda\in\Lambda}$.

We define two families of hyperbolic wavelet bases with discretization parameter $L \in \mathbb{Z}$ by the index sets

$$\Lambda_L^{(3,n)} := \left\{\lambda \in (\nabla^{(3)})^n \colon |\lambda_1| + \ldots + |\lambda_n| \le L\right\}, \quad \Lambda_L^{(1,3n)} := \left\{\lambda \in \nabla^{3n} \colon |\lambda_1| + \ldots + |\lambda_{3n}| \le L\right\}, \tag{4.3}$$

which are nonempty for $L \ge nj_0$ and $L \ge 3nj_0$, respectively. 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 (3.3), which is not the best possible, but is included here rather for illustrative purposes; a more detailed analysis can be found in [51,53].

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

In the following, we restrict the approximation to a subset Λ of indices of the hyperbolic wavelet bases $\Lambda_L^{(3,2)}$ or $\Lambda_L^{(1,6)}$ 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. *Let $u \in \mathcal{H}^{s,1}(\mathbb{R}^3;2)$, $0 < s \le 1$, satisfy a decay condition (2.6) with some $\delta > 0$, and let either $d = 3$, $D = 2$ or $d = 1$, $D = 6$, then there exists a $C > 0$ depending on s, δ, ψ and j_0 such that for each $L \ge Dj_0$ there exists a subset $\Lambda \subset \Lambda_L^{(d,D)}$ with*

$$\inf_{v \in \operatorname{span}\{\Psi_\lambda :\, \lambda \in \Lambda\}} \|u - v\|_1 \le C 2^{-\frac{d}{3}sL} \|u\|_{s,1}, \quad \text{where } \#\Lambda \lesssim L^{5+D} 2^{dL}.$$

Proof. For the following, let $\bar\Lambda_L := \Lambda_L^{(d,D)}$. Fix an $M > 0$; for $R > M$, let $\eta_R \in \mathrm{C}^\infty(\mathbb{R}^6)$, $0 \le \eta_R \le 1$, such that $\eta_R \equiv 0$ on $B_{R-M}(0)$, $\eta_R \equiv 1$ on $\mathbb{R}^6 \setminus B_R(0)$ and $\|\mathrm{D}\eta_R\|_\infty$ is bounded uniformly in R. Then with a $c > 0$ independent of R and $v \in \mathrm{H}^1(\mathbb{R}^6)$, $\|\eta_R v\|_1 \le c \|v \,|\, \mathrm{H}^1(\mathbb{R}^6 \setminus B_{R-M}(0))\|$.

Furthermore, let $\Lambda_R = \{\lambda \in \bar\Lambda_L :\, \operatorname{supp}\Psi_\lambda \cap B_R(0) \ne \emptyset\}$. Then

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

Using the wavelet characterization (4.1),

$$\|u - u_{\Lambda_R}\|_1^2 \lesssim \sum_{\lambda \notin \bar\Lambda_L} \sum_{i=1}^D 2^{2|\lambda_i|} |\langle u, \Psi_\lambda\rangle|^2 + \sum_{\lambda \notin \Lambda_R} \sum_{i=1}^D 2^{2|\lambda_i|} |\langle u, \Psi_\lambda\rangle|^2.$$

Now on the one hand, by (4.2)

$$\sum_{\lambda \notin \bar\Lambda_L} \sum_{i=1}^D 2^{2|\lambda_i|} |\langle u, \Psi_\lambda\rangle|^2 \le 2^{-2\frac{d}{3}s(L+1)} \sum_{\lambda \notin \bar\Lambda_L} 2^{2\frac{d}{3}s\sum_{i=1}^D |\lambda_i|} \sum_{i=1}^D 2^{2|\lambda_i|} |\langle u, \Psi_\lambda\rangle|^2 \lesssim 2^{-2\frac{d}{3}sL} \|u\|_{s,1}^2,$$

where we have used $\prod_{k=1}^3 2^{\frac{2ds}{3} j_k} \le \frac{1}{3} \sum_{k=1}^3 2^{2dsj_k}$ for $j \in \mathbb{Z}^3$ in the case $d = 1$, $D = 6$. On the other hand, since $u = \eta_R u$ outside $B_R(0)$,

$$\begin{aligned}\sum_{\lambda \notin \Lambda_R} \sum_{i=1}^D 2^{2|\lambda_i|} |\langle u, \Psi_\lambda\rangle|^2 &\le \sum_{\lambda} \sum_{i=1}^D 2^{2|\lambda_i|} |\langle \eta_R u, \Psi_\lambda\rangle|^2 \lesssim \|\eta_R u\|_1^2 \\ &\le c^2 \mathrm{e}^{-2\delta(R-M)} \int \mathrm{e}^{2\delta|(x,y)|} \left(|u|^2 + |\mathrm{D}u|^2\right) \mathrm{d}(x,y) \lesssim \mathrm{e}^{-2\delta R}.\end{aligned}$$

In summary, $\|u - u_{\Lambda_R}\|_1 \lesssim \mathrm{e}^{-\delta R} + 2^{-s\frac{d}{3}L} \|u\|_{s,1}$, 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 $|\lambda| = j \in \mathbb{Z}^D$, $j \ge j_0$ componentwise, is then proportional to $\prod_{i=1}^D (\max\{L2^{j_i}, 1\})^d$, and summation over $j \in \mathbb{Z}^D$ with $\sum_{i=1}^D j_i \le L$, $j \ge j_0$ yields $\#\Lambda \lesssim L^{5+D} 2^{dL}$. □

Remark 4.2. Although a deeper analysis allows to remove powers of L from $\#\Lambda$, which has been shown in [53] using exponential decay of mixed derivatives, for $s = 1$ as for the explicitly correlated formulation (3.3) the above already gives a rate of almost $(\#\Lambda)^{-1/3}$.

For comparison, we can apply Theorem 4.1 to a direct hyperbolic wavelet discretization of the standard formulation (3.2): it has been shown in [48] that the corresponding eigenfunctions are in $\mathcal{H}^{1/2,1}(\mathbb{R}^3;2)$, and one

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

obtains a convergence rate of almost $(\#\Lambda)^{-1/6}$. This regularity result has recently been sharpened to $\mathcal{H}^{s,1}(\mathbb{R}^3;2)$ for $s < \frac{3}{4}$ in [37], corresponding to almost $(\#\Lambda)^{-1/4}$.

The regularity result for the explicitly correlated formulation cannot be expected to be sharp. Note furthermore that for the convergence of eigenvalues in the nonsymmetric modified problem, on the basis of the available results we can only expect a rate of $(\#\Lambda)^{-7/12}$ due to the lower regularity of adjoint eigenfunctions, *cf.* Remark 3.4.

Remark 4.3. Both choices $\Lambda_L^{(3,2)}$ and $\Lambda_L^{(1,6)}$ yield essentially the same convergence rate, with $\Lambda_L^{(1,6)}$ requiring more unknowns by a factor proportional to L^4. Although this seems to indicate a disadvantage of the fully anisotropic basis, it should be noted that this reflects the underlying regularity assumptions, and that the comparison may change when more general sets of wavelet indices corresponding to assumptions on u different from those in Theorem 4.1 are used.

An advantage of the fully anisotropic construction that has been pointed out in [51] is that by the imbedding $\mathcal{H}^{s,1}(\mathbb{R}^3;n) \hookrightarrow \mathcal{H}^{s/3,1}(\mathbb{R};3n)$, univariate wavelets of a certain Sobolev regularity characterize spaces of mixed smoothness corresponding to a larger range of convergence rates.

However, some particularly crucial differences between the fully and partially anisotropic constructions will become apparent later in the discussion of algorithmic issues.

4.2. Suitable wavelet bases

For obtaining approximations as in Theorem 4.1 as solutions of 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, as noted in [14], 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 space dimension, *i.e.*, 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 orthogonality is that the discretization of the identity is diagonal, and 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 difficult to enforce for non-orthonormal basis functions.

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.

Taking these rather restrictive requirements together, there are not many known suitable choices. One are Daubechies wavelets [13]; a possible alternative would be the piecewise polynomial continuous multiwavelets of [15,16] as used for instance in [14].

Although polynomial multiwavelets have some advantages over Daubechies wavelets, such as better properties for matrix compression in the context of adaptive schemes and substantially easier handling of integration, there is also a major disadvantage: with increasing approximation orders and smoothness requirements, the number of different multiwavelet basis functions in the one-dimensional construction increases fairly quickly, and this number in turn enters exponentially with respect to space dimension in the number of basis functions per dyadic grid point that need to be considered in the higher-dimensional case. In the present context, this presents a serious practical obstruction, and hence the following exposition, though in principle more generally applicable, pays particular attention to Daubechies wavelets.

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

Note that least asymmetric Daubechies wavelets have also been employed for density functional theory problems with smooth solutions in 3D [22]. Previous wavelet-based discretizations of the higher-dimensional Schrödinger equation have been based *e.g.* on globally supported Meyer wavelets [24], approximately orthogonal Gaussian frames [25, 28] and semiorthogonal piecewise linear prewavelets [51].

5. Exploiting structural properties for applying operators

Up to this point, we were dealing mainly 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. Although our main interest is the modified problem (2.4), the results of this section apply to the standard formulation (2.1) as well.

As a basic option, one can consider a wavelet-Galerkin discretization of the modified eigenvalue problem (2.4) defined by a subset of wavelet indices Λ, 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 $\Lambda\subset\Lambda_L^{(3,2)}$ or $\Lambda\subset\Lambda_L^{(1,6)}$ fixed a priori corresponds to the regularity and approximation results of the previous sections. However, the concepts discussed in the following may also be used in the framework of adaptive wavelet schemes, where a basic ingredient is the approximate evaluation of

$$\left(\sum_{\mu\in\Lambda} \tilde{a}(\Psi_\mu,\Psi_\nu)\, u_\mu\right)_{\nu\in\Lambda'} \qquad \text{for some } \Lambda'\supseteq\Lambda, \tag{5.2}$$

where Λ,Λ' may be more general index sets.

Due to the anisotropy of basis functions, Galerkin discretization matrices arising from hyperbolic wavelet bases are in general almost dense. Therefore, assembling the matrices in (5.1) and (5.2) explicitly is out of the question. Instead, we aim to construct algorithms for applying these matrices to vectors with reasonable complexity. Note that by (4.1), a simple diagonal preconditioning yields condition numbers that are bounded independently of Λ,Λ'.

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), *i.e.*, the matrix entries read

$$\tilde{a}(\Psi_\mu,\Psi_\nu) = \int_{\mathbb{R}^6} \frac{1}{2}(\mathrm{D}\Psi_\mu\cdot\mathrm{D}\Psi_\nu) - \left(\frac{2}{|x|}+\frac{2}{|y|}\right)\Psi_\mu\Psi_\nu - \frac{1}{2}\frac{x-y}{|x-y|}\cdot(\mathrm{D}_x-\mathrm{D}_y)\Psi_\mu\Psi_\nu\,\mathrm{d}(x,y) - \frac{1}{4}\delta_{\mu\nu}. \tag{5.3}$$

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 basic algorithms outlined in the following subsection. Especially in view of the two-electron operator, such a scheme by itself is not sufficient for (5.3). One could proceed by directly using matrix compression based on the vanishing moment properties of wavelets, resulting in a perturbed application of operators.

However, even assembling only parts required in the compressed application of the matrices corresponding to one- and two-electron potential operators in (5.3) is too expensive in practice: the number of nonzero entries per column in these matrices generally scales like the third or sixth power, respectively, of the support size of the scaling functions, which, particularly for Daubechies wavelets, is quite large.

This leads us to consider in Section 5.2 approximations for the coefficients that introduce auxiliary tensor product structures, which can also be seen as a very problem-specific type of compression. Making algorithmic use of such separable structures as described in the following subsection, the approximate operators can be

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

applied with a scaling that is linear or quadratic, respectively, in the scaling function support size. The numerical work required for applying (5.3), both in terms of storage and of operators, is thus reduced by several orders of magnitude.

5.1. Matrix-vector products for hyperbolic wavelet discretizations

For operators that have a tensor product structure, efficient schemes for applying discretization matrices arising from subsets of $(\nabla^{(d)})^D$, in particular of $\Lambda_L^{(d,D)}$, can be obtained by exploiting this product structure for certain blocks corresponding to different combinations of levels. To illustrate the schemes, we first consider a tensor product operator $A = A^{(1)} \otimes A^{(2)}$ and for some $\Lambda, \Lambda' \subset (\nabla^{(d)})^2$, $d \in \mathbb{N}$, the task of applying the section

$$A_{\Lambda',\Lambda} = (\langle A\Psi_\mu, \Psi_\nu\rangle)_{\nu\in\Lambda',\mu\in\Lambda}\,, \quad \langle A\Psi_\mu, \Psi_\nu\rangle = \langle A^{(1)}\Psi_{\mu_1}, \Psi_{\nu_1}\rangle\langle A^{(2)}\Psi_{\mu_2}, \Psi_{\nu_2}\rangle$$

to a vector $u = (u_\mu)_{\mu\in\Lambda}$. Here one is facing the difficulty that for Λ corresponding to a sparse grid, $A_{\Lambda',\Lambda}$ itself does not have Kronecker product structure.

The basis of the scheme introduced by Schwab and Todor in [43] for applying $A_{\Lambda',\Lambda}$ to a vector are partitions

$$\Lambda = \bigcup_{i\in\mathcal{I}} \Lambda_i, \quad \Lambda' = \bigcup_{j\in\mathcal{I}'} \Lambda'_j \qquad \text{with } \Lambda_i = \Lambda_{i,1}\times\Lambda_{i,2},\ \ \Lambda'_j = \Lambda'_{j,1}\times\Lambda'_{j,2}. \tag{5.4}$$

If such a decomposition holds, $A_{\Lambda'_j,\Lambda_i} = A^{(1)}_{\Lambda'_{j,1},\Lambda_{i,1}} \otimes A^{(2)}_{\Lambda'_{j,2},\Lambda_{i,2}}$ for $i \in \mathcal{I}, j \in \mathcal{I}'$, and furthermore

$$A_{\Lambda'_j,\Lambda_i} = \begin{cases} (A^{(1)}_{\Lambda'_{j,1},\Lambda_{i,1}} \otimes \mathrm{I}_{\Lambda'_{j,2}})(\mathrm{I}_{\Lambda_{i,1}} \otimes A^{(2)}_{\Lambda'_{j,2},\Lambda_{i,2}}) & \text{(a)},\\ (\mathrm{I}_{\Lambda'_{j,1}} \otimes A^{(2)}_{\Lambda'_{j,2},\Lambda_{i,2}})(A^{(1)}_{\Lambda'_{j,1},\Lambda_{i,1}} \otimes \mathrm{I}_{\Lambda_{i,2}}) & \text{(b)}, \end{cases} \tag{5.5}$$

where $\mathrm{I}_{\tilde\Lambda}$ denotes the identity on each index set $\tilde\Lambda$. Now for each section corresponding to a pairing $(i,j) \in \mathcal{I}\times\mathcal{I}'$, one can pick the order of applying the two factors in (5.5) that minimizes the size of the resulting intermediates.

Consider as an example a uniformly refined sparse grid, with $\Lambda = \Lambda' \subset \Lambda_L^{(d,2)}$, assuming that for each combination of wavelet levels the sets of spatial indices have product structure. Let $\mathcal{I} = \{i \in \mathbb{Z}^2 \colon i \geq j_0,\ i_1+i_2 \leq L\}$, and let $\Lambda_i = \{\lambda \in \Lambda \colon |\lambda_1| = i_1, |\lambda_2| = i_2\}$ for $i \in \mathcal{I}$, where uniform refinement yields $\#\Lambda_{i,k} \sim 2^{di_k}$ and $\#\Lambda \sim L2^{dL}$. Then case (a) in (5.5) is used if $i_1 + j_2 \leq j_1 + i_2$, and case (b) otherwise. If A is a local operator, it can be shown that the operation count for applying $A_{\Lambda',\Lambda}$ in this manner is proportional to $L^3 2^{dL}$, *cf.* [52]; moreover, the scheme can be recursively extended to higher-dimensional sparse grids, *cf.* [29].

The above method can be applied to fairly general subsets Λ of the full tensor product basis and stores only minimal intermediate results. It is, however, desirable to further reduce the number of required operations.

For the description of a scheme that can reduce the number of required operations at the expense of storing larger intermediate vectors, we impose more specific assumptions than (5.4) on Λ, Λ' from the outset: let $\mathcal{I}, \mathcal{I}' \subset \{i \in \mathbb{Z}^2 \colon i \geq j_0\}$ and let (5.4) hold with $\Lambda_i = \{\lambda \in \Lambda \colon |\lambda_1| = i_1, |\lambda_2| = i_2\} = \Lambda_{i,1}\times\Lambda_{i,2}$ for $i \in \mathcal{I}$, $\Lambda'_j = \{\lambda \in \Lambda' \colon |\lambda_1| = j_1, |\lambda_2| = j_2\} = \Lambda'_{j,1}\times\Lambda'_{j,2}$ for $j \in \mathcal{I}'$.

Furthermore, we need the notations $\mathcal{I}_k = \{i_k \colon i \in \mathcal{I}\}$, $\mathcal{I}'_k = \{i_k \colon i \in \mathcal{I}'\}$ for $k = 1, 2$, and $\mathcal{I}_1(i_2) = \{i_1 \colon (i_1, i_2) \in \mathcal{I}\}$, with $\mathcal{I}_2(i_1), \mathcal{I}'_1(i_2), \mathcal{I}'_2(i_1)$ defined analogously. For $k = 1, 2$, let $N_k \in \mathbb{N}$ and $\mathcal{N}_k = \{1, \ldots, N_k\}$; for each $j_k \in \mathcal{I}'_k$, choose $\mathcal{J}_{k,n}(j_k) \subseteq \mathcal{I}_k$, $n \in \mathcal{N}_k$, to be a partition of $\mathcal{I}_k$, that is, $\mathcal{I}_k = \bigcup_{n\in\mathcal{N}_k} \mathcal{J}_{k,n}(j_k)$ for each $j_k \in \mathcal{I}'_k$.

We then have $A_{\Lambda',\Lambda}u = \sum_{n\in\mathcal{N}_1\times\mathcal{N}_2} v_{\Lambda',n}$, where for each $j \in \mathcal{I}'$,

$$v_{\Lambda'_j,n} = \begin{cases} \displaystyle\sum_{\substack{i_1\in\mathcal{I}_1\cap\\ \mathcal{J}_{1,n_1}(j_1)}} (A^{(1)}_{\Lambda'_{j_1},\Lambda_{i_1}} \otimes \mathrm{I}) \sum_{\substack{i_2\in\mathcal{I}_2(i_1)\cap\\ \mathcal{J}_{2,n_2}(j_2)}} (\mathrm{I} \otimes A^{(2)}_{\Lambda'_{j_2},\Lambda_{i_2}})\, u_{\Lambda_i},\\ \displaystyle\sum_{\substack{i_2\in\mathcal{I}_2\cap\\ \mathcal{J}_{2,n_2}(j_2)}} (\mathrm{I} \otimes A^{(2)}_{\Lambda'_{j_2},\Lambda_{i_2}}) \sum_{\substack{i_1\in\mathcal{I}_1(i_2)\cap\\ \mathcal{J}_{1,n_1}(j_1)}} (A^{(1)}_{\Lambda'_{j_1},\Lambda_{i_1}} \otimes \mathrm{I})\, u_{\Lambda_i}. \end{cases} \tag{5.6}$$

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

The inner sums in (5.6) can be evaluated first, with results stored in an intermediate vector, where for each (n_1, n_2), one chooses the order of summation in (5.6) that minimizes number of operations and intermediate storage.

When $N_1 = N_2 = 1$ and $\mathcal{J}_{1,1}(j_1) = \mathcal{I}_1$, $\mathcal{J}_{2,1}(j_2) = \mathcal{I}_2$ for all j, the intermediate results will in general have the complexity of a full tensor product grid; if the $\mathcal{J}_{k,n}$ are chosen to be the one-element subsets of $\mathcal{I}_k$, one obtains the scheme of Schwab and Todor as in (5.5).

In the case of the above example of a uniformly refined sparse grid with $\Lambda = \Lambda'$ and $\#\Lambda \sim L2^{dL}$, one obtains with $N_1 = 1$, $\mathcal{J}_{1,1}(j_1) = \mathcal{I}_1$ and $N_2 = 2$, $\mathcal{J}_{2,1}(j_2) = \{i_2 \in \mathcal{I}_2 \colon\ i_2 > j_2\}$, $\mathcal{J}_{2,2}(j_2) = \{i_2 \in \mathcal{I}_2 \colon\ i_2 \le j_2\}$ a splitting of the matrix-vector product as

$$A_{\Lambda_j,\Lambda} u = \sum_{i_1 \in \mathcal{I}_1} (A^{(1)}_{\Lambda_{j_1},\Lambda_{i_1}} \otimes \mathrm{I}) \sum_{\substack{i_2 \in \mathcal{I}_2(i_1) \\ j_2 < i_2}} (\mathrm{I} \otimes A^{(2)}_{\Lambda_{j_2},\Lambda_{i_2}})\, u_{\Lambda_i} + \sum_{\substack{i_2 \in \mathcal{I}_2 \\ j_2 \ge i_2}} (\mathrm{I} \otimes A^{(2)}_{\Lambda_{j_2},\Lambda_{i_2}}) \sum_{i_1 \in \mathcal{I}_1(i_2)} (A^{(1)}_{\Lambda_{j_1},\Lambda_{i_1}} \otimes \mathrm{I})\, u_{\Lambda_i}, \tag{5.7}$$

where the intermediate values arising from the evaluation of the inner sums can be stored, simultaneously for all j_2 and j_1, respectively, with a complexity proportional to $\#\Lambda$. This is known as the unidirectional scheme [4,8]. In this example, the number of operations for computing the matrix-vector product can be shown to be of order $L^2 2^{dL}$, *cf.* [52].

For more general index sets Λ, Λ' it can of course be advantageous to select sets $\mathcal{J}_{k,n}$ different from the above examples. Again, the scheme based on (5.6) is straightforward to extend to higher dimensions.

We describe the application of the above schemes defined by (5.5) and (5.6) to our setting in more detail in Section 5.2.2, after introducing the separable approximations of potential terms that they require.

5.2. Separable approximation of potentials

The basic idea for reducing the complexity of applying one- and two-electron potential operators in (5.3) is to replace them by separable approximations, in the present two-electron case

$$\frac{1}{|x|} \approx \sum_{k=1}^{N_f} f_k(x_1)\, f_k(x_2)\, f_k(x_3), \qquad \frac{x-y}{|x-y|} \approx (x-y) \sum_{k=1}^{N_g} g_k(x_1-y_1)\, g_k(x_2-y_2)\, g_k(x_3-y_3). \tag{5.8}$$

In what follows, we will use approximations of the same type for f_k and g_k, although different constructions are possible. In the particular cases in (5.8), we take $f_k = \omega_k^{1/3} \exp(-\alpha_k |\cdot|^2)$ with suitable $\omega_k, \alpha_k > 0$. This enables surprisingly efficient approximations, see *e.g.* (5.18) as obtained in [7]. In the context of electronic structure calculations, such expansions have also been used for the purpose of separable approximations of Green's functions, see *e.g.* [21,30].

Note that when correlation factors are required to remain bounded, for approximations as in (5.8) to remain applicable, one may modify (2.2) by a suitable function that is itself separable, *e.g.* a Gaussian damping factor.

To simplify notation for the discussion of discretization matrices arising when using (5.8), for $\lambda \in \nabla^{(d)}$ let $p_i(\lambda) := (|\lambda|, \mathrm{k}_i(\lambda), \mathrm{s}_i(\lambda))$, $i = 1, \dots, d$. For the Coulomb potential on $\mathbb{R}^3$, if $\mu, \nu \in \nabla^{(3)}$, we obtain

$$\int_{\mathbb{R}^3} \frac{1}{|x|} \Psi_\mu \Psi_\nu \,\mathrm{d}x \approx \sum_k \prod_{i=1}^{3} \int_{\mathbb{R}} f_k\, \psi_{p_i(\mu)} \psi_{p_i(\nu)} \,\mathrm{d}x, \tag{5.9}$$

whereas if $\mu, \nu \in \nabla^3$,

$$\int_{\mathbb{R}^3} \frac{1}{|x|} \Psi_\mu \Psi_\nu \,\mathrm{d}x \approx \sum_k \prod_{i=1}^{3} \int_{\mathbb{R}} f_k\, \psi_{\mu_i} \psi_{\nu_i} \,\mathrm{d}x; \tag{5.10}$$

for the Coulomb potential on $\mathbb{R}^6$, if $\mu, \nu \in (\nabla^{(3)})^2$,

$$\int_{\mathbb{R}^6} \frac{1}{|x-y|} \Psi_\mu \Psi_\nu \,\mathrm{d}(x,y) \approx \sum_k \prod_{i=1}^{3} \int_{\mathbb{R}^2} f_k(x-y)\, \psi_{p_i(\mu_1)}(x)\, \psi_{p_i(\mu_2)}(y)\, \psi_{p_i(\nu_1)}(x)\, \psi_{p_i(\nu_2)}(y) \,\mathrm{d}(x,y), \tag{5.11}$$

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

and if $\mu, \nu \in \nabla^6$,

$$\int_{\mathbb{R}^6} \frac{1}{|x-y|} \Psi_\mu \Psi_\nu \,\mathrm{d}(x,y) \approx \sum_k \prod_{i=1}^{3} \int_{\mathbb{R}^2} f_k(x-y)\, \psi_{\mu_i}(x)\, \psi_{\nu_i}(x)\, \psi_{\mu_{i+3}}(y)\, \psi_{\nu_{i+3}}(y) \,\mathrm{d}(x,y); \tag{5.12}$$

and for the modified potential (3.6), if $\mu, \nu \in \nabla^6$,

$$\begin{aligned} &\int_{\mathbb{R}^6} \frac{x-y}{|x-y|} \cdot (\mathrm{D}_x - \mathrm{D}_y) \Psi_\mu \, \Psi_\nu \,\mathrm{d}(x,y) \\ &\quad \approx \sum_k \sum_j \int_{\mathbb{R}^2} (x-y) g_k(x-y) \,(\mathrm{D}_x - \mathrm{D}_y) \left(\psi_{\mu_j}(x)\, \psi_{\mu_{j+3}}(y)\right)\, \psi_{\nu_j}(x)\, \psi_{\nu_{j+3}}(y) \,\mathrm{d}(x,y) \\ &\qquad \times \prod_{i \neq j} \int_{\mathbb{R}^2} g_k(x-y)\, \psi_{\mu_i}(x)\, \psi_{\mu_{i+3}}(y)\, \psi_{\nu_i}(x)\, \psi_{\nu_{i+3}}(y) \,\mathrm{d}(x,y), \end{aligned} \tag{5.13}$$

and analogously, with the same modifications as between (5.11) and (5.12), in the case $\mu, \nu \in (\nabla^{(3)})^2$. Note that in the case of indices in ∇^6 as in (5.10), (5.12), (5.13), the factors on the right hand side are again discretization matrices corresponding to ∇ or ∇^2, whereas this is not the case for indices in $(\nabla^{(3)})^2$ as in (5.9), (5.11).

In connection with wavelet bases, in [10,19,38] an approach for efficient computation of integrals using separable approximations has been developed, focusing on the efficient computation of individual wavelet coefficients and discretization matrix entries and accordingly tailored error estimates. Since for our purposes we want to avoid explicitly assembling matrices as in (5.3) whenever possible, we are rather interested in a different point of view: we replace the potential terms in (5.3) by separable approximations (5.8) 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.

The same approach to the application of operators was taken in the recent work [51] for the Coulomb potentials in the original formulation (3.2) of the electronic Schrödinger equation, and similar estimates for the energy error have been obtained. The convergence theory given in this section provides an extension to the modified problem (3.3), additionally provides estimates for the error in eigenfunctions and yields qualitatively better convergence of the energy if the wavelets are sufficiently smooth.

Before turning to the error analysis, we briefly discuss the algorithmic use of expansions of the form (5.8).

5.2.1. 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.* [5,12]); in particular, this suffices for the computation of $\int_{\mathbb{R}} \psi'_\mu \psi'_\nu \,\mathrm{d}x$, $\mu, \nu \in \nabla$, as required for the Laplacian terms.

When using the approximations (5.8), it is also necessary to integrate products of Gaussians with wavelets as in equations (5.9) to (5.13). This can be done using an auxiliary wavelet family where the scaling function has vanishing moments, *i.e.*, satisfies $\int x^k \varphi \,\mathrm{d}x = 0$ for some range $k = 1, \ldots, K$, and hence provides an efficient one-point integration rule, for example the Deslaurier-Dubuc-Sweldens wavelets [45] as suggested in [17]. This procedure is reasonably efficient when applied to the one-electron case as in (5.9) and (5.10), but becomes very expensive for large values of α – which are required in our context – in the case of two-electron terms as in (5.11), (5.12), and (5.13).

An integration scheme proposed in [40] has been employed successfully for computations involving smooth pseudopotentials in 3D with Daubechies wavelets, but is not suitable for our purposes due to its smoothness requirements on both potentials and wavefunctions.

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

An alternative scheme can be based on the observation that, for $\mu, \nu \in \nabla^2$,

$$\int_{\mathbb{R}^2} \mathrm{e}^{-\alpha(x-y)^2}\, \Psi_\nu\, \Psi_\mu \,\mathrm{d}(x,y) = \frac{1}{2\sqrt{\alpha\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\xi^2/(4\alpha)} \overline{(\psi_{\nu_1}\, \psi_{\mu_1})^\wedge(\xi)}\; (\psi_{\nu_2}\, \psi_{\mu_2})^\wedge(\xi)\,\mathrm{d}\xi. \tag{5.14}$$

Note that a similar identity was used in [24] for two-electron integrals with full potentials, whose Fourier transforms are singular in general, and (globally supported) Meyer wavelets, which have an explicit representation in Fourier domain.

When instead using compactly supported wavelets such as the Daubechies family combined with the separable approximations by Gaussians, the integrand on the right hand side of (5.14) is an analytic function. Hence an exponentially convergent quadrature scheme can be constructed on the basis of the trapezoidal rule; a prerequisite is a method for evaluating Fourier transforms of products of wavelets. Details of this approach will be given in a forthcoming publication [3].

5.2.2. Matrix-vector products

In order to take advantage of the separable structure of the approximate potentials, we may employ the algorithms outlined in Subsection 5.1 to operate along directions separately. In addition, we can take advantage of vanishing moments of the wavelets to compress the involved lower-dimensional operators.

To enable effective use of the separable structure, the sets of wavelet indices Λ being operated on need to have certain tensor structures themselves. For atomic systems, for instance, it is reasonable to impose the simple restriction that the set of translation parameters $\{\lambda \in \Lambda\colon\ |\lambda| = j,\, \mathrm{s}(\lambda) = s\}$ for each j, s can be written as a cartesian product.

Under this assumption, the general variant of the unidirectional scheme given by (5.6), adapted to operators given as a Kronecker product of three factors, can be applied to the summands in (5.10), (5.12), and (5.13) corresponding to the basis $\{\Psi_\lambda\colon\ \lambda \in \nabla^6\}$. The scheme in (5.6) and (5.7) is not applicable, however, to approximations of operators as in (5.11) corresponding to the basis $\{\Psi_\lambda\colon\ \lambda \in (\nabla^{(3)})^2\}$, since one does not have the necessary tensor structure for the index sets corresponding to the factors; in this case, one can still use the algorithm of Schwab and Todor as in (5.5).

We now turn to the compression by dropping small entries of factor matrices as in (5.12), (5.13) that arise for the basis $\{\Psi_\lambda\colon\ \lambda \in \nabla^6\}$. As will become clear, such techniques are of limited use in the case of $\{\Psi_\lambda\colon\ \lambda \in (\nabla^{(3)})^2\}$.

From the definition of Sobolev norms in terms of the Fourier transform as in (3.1), one immediately obtains the following, which will be used again later. We refer to [51] for a proof, see also [26,27] for the case of bounded domains.

Lemma 5.1. *Let $s > 0$, $n_i \in \mathbb{N}$ for $i = 1, \dots, N$. Then*

$$\mathrm{H}^s(\mathbb{R}^{\sum_i n_i}) = \mathrm{H}^s(\mathbb{R}^{n_1}) \otimes \mathrm{L}^2(\mathbb{R}^{\sum_{i>1} n_i}) \cap \dots \cap \mathrm{L}^2(\mathbb{R}^{\sum_{i<j} n_i}) \otimes \mathrm{H}^s(\mathbb{R}^{n_j}) \otimes \mathrm{L}^2(\mathbb{R}^{\sum_{i>j} n_i}) \cap \dots \cap \mathrm{L}^2(\mathbb{R}^{\sum_{i<N} n_i}) \otimes \mathrm{H}^s(\mathbb{R}^{n_N}).$$

For the compression of operators with tensor structure that we are interested in, Lemma 5.1 has the implication that for a given operator $M\colon \mathrm{H}^1(\mathbb{R}^{dD}) \to \mathrm{H}^{-1}(\mathbb{R}^{dD})$ with $M = M^{(1)} \otimes \dots \otimes M^{(D)}$, where $M^{(k)}\colon \mathrm{L}^2(\mathbb{R}^d) \to \mathrm{L}^2(\mathbb{R}^d)$, $k = 1, \dots, D$, and for a compressed approximation $\tilde{M} = \tilde{M}^{(1)} \otimes \dots \otimes \tilde{M}^{(D)}$ obtained by dropping certain entries from the wavelet representation of each $M^{(k)}$, we have

$$\|M - \tilde{M}\|_{\mathrm{H}^1 \to \mathrm{H}^{-1}} \le C^{D-1} \sum_{i=1}^{D} \|M^{(i)} - \tilde{M}^{(i)}\|_{\mathrm{H}^1 \to \mathrm{H}^{-1}}, \tag{5.15}$$

where $C = \max_i\{\|M^{(i)}\|_{\mathrm{L}^2 \to \mathrm{L}^2}, \|\tilde{M}^{(i)}\|_{\mathrm{L}^2 \to \mathrm{L}^2}\}$. The compression of tensor product operators is therefore reduced to the compression of their factors, where (5.15) applies with $d = 1$, $D = 3$ to the summands in (5.10) and with $d = 2$, $D = 3$ to those in (5.12). The situation is slightly different for (5.13), since there one of the factors is

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

not bounded as an operator on $\mathrm{L}^2(\mathbb{R}^2)$. However, if *e.g.* the first factor $M^{(1)}$ is only assumed to be bounded $\mathrm{H}^1(\mathbb{R}^d) \to \mathrm{L}^2(\mathbb{R}^d)$, one still has

$$\|M - \tilde{M}\|_{\mathrm{H}^1\to\mathrm{H}^{-1}} \leq C^{D-1}\left(\|M^{(1)} - \tilde{M}^{(1)}\|_{\mathrm{H}^1\to\mathrm{L}^2} + \sum_{i=2}^{D}\|M^{(i)} - \tilde{M}^{(i)}\|_{\mathrm{L}^2\to\mathrm{H}^{-1}}\right),$$

where $C = \max\{\|M^{(1)}\|_{\mathrm{H}^1\to\mathrm{L}^2}, \|\tilde{M}^{(1)}\|_{\mathrm{H}^1\to\mathrm{L}^2}, \max_{i>1}\|M^{(i)}\|_{\mathrm{L}^2\to\mathrm{H}^{-1}}, \max_{i>1}\|\tilde{M}^{(i)}\|_{\mathrm{L}^2\to\mathrm{H}^{-1}}\}$, which can be applied to the summands in (5.13).

For a discussion of compressibility properties of the Laplacian, see *e.g.* [11,44]; better results in this direction are known for piecewise polynomial (multi)wavelets than for Daubechies wavelets, *cf.* [44]. We next give an estimate that can serve as the basis of similar compression schemes for the factors in (5.12), that is, for matrices $(m_{\nu\mu})$ with entries

$$m_{\nu\mu} = \int_{\mathbb{R}^2} \mathrm{e}^{-\alpha(x-y)^2}\psi_{\mu_1}(x)\,\psi_{\mu_2}(y)\,\psi_{\nu_1}(x)\,\psi_{\nu_2}(y)\,\mathrm{d}(x,y), \quad \nu,\mu\in\nabla^2.$$

These matrices also appear in the case of the nonsymmetric modified potential, where analogous arguments can be applied to the additional factor matrices involving derivatives.

Assuming that ψ has at least p vanishing moments and that $\psi \in C^p(\mathbb{R})$, for $m_1, m_2 \in \mathbb{N}$ with $m_1, m_2 \leq p$, if *e.g.* $|\nu_1| \geq |\mu_1|$ and $|\mu_2| \geq |\nu_2|$, we have

$$|m_{\nu\mu}| \lesssim \|\psi_{\nu_1}\otimes\psi_{\mu_2}\|_{\mathrm{L}^\infty}\, 2^{-m_1|\nu_1|-m_2|\mu_2|}\int_{\operatorname{supp}\psi_{\nu_1}\otimes\psi_{\mu_2}}\left|\mathrm{D}_x^{m_1}\mathrm{D}_y^{m_2}\left(\mathrm{e}^{-\alpha(x-y)^2}\psi_{\mu_1}(x)\psi_{\nu_2}(y)\right)\right|\mathrm{d}(x,y), \tag{5.16}$$

and analogously for the other possible combinations of levels. Note that in general, such an estimate is not possible for the factors in (5.11) corresponding to $(\nabla^{(3)})^2$, since in that case one also has to deal with products of scaling functions on higher levels. Making use of Cramérs inequality [31] for the derivatives of the Gaussian term in (5.16), for each $(j_1, j_2) \in \mathbb{Z}^2$ with $j_1, j_2 \geq j_0$ we obtain

$$\begin{aligned}\sum_{\substack{\mu\in\nabla^2\\ |\mu_k|=j_k}} |m_{\nu\mu}| \;&\lesssim 2^{-\max_k|\nu_k|}\min\left\{\alpha^{-\frac{1}{2}}, 2^{-\min_k|\nu_k|}\sup_{(x,y)\in\operatorname{supp}\Psi_\nu}\mathrm{e}^{-\frac{\alpha}{2}(x-y)^2}\right\}\\ &\quad\times\prod_{k=1}^{2}2^{\frac{1}{2}(|\nu_k|+j_k)}\left(\min\left\{1, 2^{-\max\{|\nu_k|,j_k\}}\max\left\{\alpha^{\frac{1}{2}}, 2^{\min\{|\nu_k|,j_k\}}\right\}\right\}\right)^p.\end{aligned}$$

In combination with Lemma 5.5 below, this estimate can serve as the basis of matrix compression schemes.

We may conclude from the above considerations that although the fully anisotropic tensor product basis $\{\Psi_\lambda\colon\ \lambda\in\nabla^6\}$ does not offer an advantage over $\{\Psi_\lambda\colon\ \lambda\in(\nabla^{(3)})^2\}$ in terms of theoretically guaranteed approximation rates, it is much better suited for computations when using separable approximations for two-electron potentials.

5.2.3. Approximation of one-electron potentials

We now give a concrete construction for separable approximations (5.8) for the one-electron Coulomb potential and estimate the error in operator norm caused by this approximation when combined with a wavelet discretization. The construction is based on an estimate for approximation on $[1, S]$, $S > 1$, from [7],

$$\left\|\frac{1}{\sqrt{t}} - \sum_{k=1}^{N}\omega_k\exp(-\alpha_k t)\right\|_{\infty,[1,S]} \leq 8\sqrt{2}\exp\left(-\frac{\pi^2 N}{\ln(8S)}\right) =: \delta(N,S), \tag{5.17}$$

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

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

$$\left\| \frac{1}{|x|} - \sum_{k=1}^{N} \frac{\omega_k}{r} \exp\left(-\frac{\alpha_k}{r^2}|x|^2\right) \right\|_{\infty,\{r\le|x|\le R\}} \le r^{-1}\delta(N, r^{-2}R^2). \tag{5.18}$$

In what follows, we use the notation $p_{N,r,R}(x) := \sum_{k=1}^{N} \frac{\omega_k}{r} \exp\left(-\frac{\alpha_k}{r^2}|x|^2\right)$. We additionally introduce the abbreviation $\mathbb{Z}_{j_0} := \{j \in \mathbb{Z} \colon j \ge j_0\}$ and, for $D \in \mathbb{N}$, the set of weighting factor matrices

$$W_D := \left\{ \tau \in \mathbb{R}^{\mathbb{Z}_{j_0}^D \times \mathbb{Z}_{j_0}^D} \colon \tau_{i,j} = \tau_{j,i},\ \tau_{i,j} > 0,\ \sup_{i \in \mathbb{Z}_{j_0}^D} \sum_{j \in \mathbb{Z}_{j_0}^D} \tau_{i,j} < \infty \right\}. \tag{5.19}$$

Theorem 5.2. *Let $0 < r < R$, $\varepsilon_0 > 0$, and either $s = 1$ or $s \in (\frac{3}{2}, 2]$. For each $N \in \mathbb{N}$, let $p_N := p_{N,r,R}$. Furthermore, let either*

(1) *$\bar{\Lambda} = \nabla^{(3)}$ and $\Lambda \subset \bar{\Lambda}$ such that $\operatorname{supp}\Psi_\nu \subset B_R(0)$ for all $\nu \in \Lambda$; let $\tau \in W_1$ and for $i, j \in \mathbb{Z}_{j_0}$, let $N_{i,j}$ be chosen such that*

$$r^{-1}\delta(N_{i,j}, r^{-2}R^2) \le 2^{s(i+j)}\tau_{i,j}\varepsilon_0; \tag{5.20}$$

or

(2) *$\bar{\Lambda} = (\nabla^{(1)})^3$ and $\Lambda \subset \bar{\Lambda}$ such that $\operatorname{supp}\Psi_\nu \subset B_R(0)$ for all $\nu \in \Lambda$; let $\tau \in W_3$ and for $\alpha, \beta \in \mathbb{Z}_{j_0}^3$, let $N_{\alpha,\beta}$ be chosen such that*

$$r^{-1}\delta(N_{\alpha,\beta}, r^{-2}R^2) \le \left(\sum_{k,l=1}^{3} 2^{2s(\alpha_k+\beta_l)} \right)^{\frac{1}{2}} \tau_{\alpha,\beta}\varepsilon_0. \tag{5.21}$$

Let $A_\Lambda, \tilde{A}_\Lambda \colon \mathrm{H}^s(\mathbb{R}^3) \to \mathrm{H}^{-s}(\mathbb{R}^3)$ be the operators defined by the matrices $(a_{\nu\mu})_{\nu,\mu\in\Lambda}$, $(\tilde{a}_{\nu\mu})_{\nu,\mu\in\Lambda}$ with

$$a_{\nu\mu} = \int_{\mathbb{R}^3} \frac{1}{|x|} \Psi_\mu \Psi_\nu \,\mathrm{d}x, \quad \tilde{a}_{\nu\mu} = \int_{\mathbb{R}^3} p_{N_{|\mu|,|\nu|}} \Psi_\mu \Psi_\nu \,\mathrm{d}x,$$

extended to $\mu, \nu \in \bar{\Lambda} \setminus \Lambda$ by zero. Then there exist $C_s, C_{\psi,j_0,s,\tau} > 0$ independent of r, R, ε_0 and Λ such that

$$\left\| A_\Lambda - \tilde{A}_\Lambda \,\middle|\, \mathrm{H}^s(\mathbb{R}^3) \to \mathrm{H}^{-s}(\mathbb{R}^3) \right\| \le C_s\, r^{\eta_s} + C_{\psi,j_0,s,\tau}\, \varepsilon_0, \tag{5.22}$$

where $\eta_1 = 1$, and $\eta_s = 2$ for $s \in (\frac{3}{2}, 2]$.

The practical implication of the conditions (5.20) and (5.21) is that the separable approximation need not be uniformly accurate, but can be coarser for combinations of wavelets on higher levels. The proof is prepared by several auxiliary results, some of which will also be used in the following subsection. To obtain the estimate (5.22), the error will be split into two components corresponding to the domains $B_r(0)$ and $B_R(0) \setminus B_r(0)$. The first part is dealt with by the following proposition; for the case $s = 1$, this was shown in [51].

Proposition 5.3. *Let either $s = 1$ or $s > \frac{3}{2}$. There exists a $C > 0$ such that for $r > 0$,*

$$\sup\left\{ \int_{B_r(0)} \frac{1}{|x|} f(x)\, g(x) \,\mathrm{d}x \colon\ f, g \in \mathrm{H}^s(\mathbb{R}^3),\ \|f\|_s = \|g\|_s = 1 \right\} \le C r^{\eta_s},$$

where $\eta_1 = 1$, and $\eta_s = 2$ for $s > \frac{3}{2}$.

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

Proof. For $s = 1$, we use the argument from [51]: by Hardy's inequality, using $r|x|^{-1} > 1$ on $B_r(0)$, for any $f, g \in \mathrm{H}^1(\mathbb{R}^3)$,

$$\left| \int_{B_r(0)} \frac{1}{|x|} f(x)\, g(x)\, \mathrm{d}x \right| \leq r \int_{B_r(0)} \frac{1}{|x|^2} |f(x)\, g(x)|\, \mathrm{d}x \leq r \||x|^{-1} f\|_{\mathrm{L}^2(\mathbb{R}^3)} \||x|^{-1} g\|_{\mathrm{L}^2(\mathbb{R}^3)} \leq Cr \|f\|_1 \|g\|_1.$$

In the case $s > \frac{3}{2}$, for any $f, g \in \mathrm{H}^s(\mathbb{R}^3)$, the claim follows by Hölder's inequality and continuity of the imbedding $\mathrm{H}^s(\mathbb{R}^3) \hookrightarrow \mathrm{L}^\infty(\mathbb{R}^3)$ according to the assumption on s (*cf.* [1], Thm. 7.34). □

The following lemma will be used in estimating the error component corresponding to $B_R(0) \setminus B_r(0)$. Note that χ_d is defined as in Section 4. The proof is an adaptation of standard arguments, *cf.* [11], Section 4.6.

Lemma 5.4. *Let $d, D \in \mathbb{N}$, $0 < s \leq 2$, $M \colon \mathrm{H}^s(\mathbb{R}^{dD}) \to \mathrm{H}^{-s}(\mathbb{R}^{dD})$ be a local operator, and $m_{\nu\mu} = \langle M\Psi_\mu, \Psi_\nu \rangle$ for $\mu, \nu \in (\nabla^{(d)})^D$. If for $\tau \in W_D$ and $\varepsilon_0 > 0$,*

$$|m_{\nu\mu}| \lesssim \varepsilon_0 \left(\sum_{i=1}^D 2^{2s|\nu_i|} \sum_{j=1}^D 2^{2s|\mu_j|} \right)^{\frac{1}{2}} 2^{-\frac{d}{2} \sum_i ||\mu_i| - |\nu_i||} \tau_{|\nu|,|\mu|} \prod_{k=1}^D \chi_d(\mu_k, \nu_k),$$

then the estimate $\|M \,|\, \mathrm{H}^s(\mathbb{R}^{dD}) \to \mathrm{H}^{-s}(\mathbb{R}^{dD})\| \lesssim \varepsilon_0$ holds, with constant depending only on d, D, s, τ, and the wavelet basis.

Recall that for $M \colon \mathrm{H}^s(\mathbb{R}^{dD}) \to \mathrm{H}^{-s}(\mathbb{R}^{dD})$, $0 < s \leq 2$, by (4.1) we have

$$\|M \,|\, \mathrm{H}^s \to \mathrm{H}^{-s}\| \sim \left\| \left(\left(\sum_{i=1}^D 2^{2s|\mu_i|} \right)^{-\frac{1}{2}} \left(\sum_{i=1}^D 2^{2s|\nu_i|} \right)^{-\frac{1}{2}} \langle M\Psi_\mu, \Psi_\nu \rangle \right)_{\mu,\nu} \right\|_{\ell^2((\nabla^{(d)})^D) \to \ell^2((\nabla^{(d)})^D)}. \tag{5.23}$$

With (5.23), the proof of Lemma 5.4 can be reduced to Schur's lemma.

Lemma 5.5 (Schur's lemma, *e.g.* [39])**.** *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 $\|M\| \leq 1$.

Proof of Lemma 5.4. Let $s_j := \left(\sum_i 2^{2sj_i} \right)^{-1/2}$ for $j \in \mathbb{Z}^D$. For suitable positive weight sequences $\{\omega_\nu\}$,

$$\omega_\nu^{-1} \sum_{\mu \in (\nabla^{(d)})^D} \omega_\mu |m_{\nu\mu}| s_{|\mu|} s_{|\nu|} \lesssim \varepsilon_0\, \omega_\nu^{-1} \sum_{\mu \in (\nabla^{(d)})^D} \omega_\mu 2^{-\frac{d}{2} \sum_i ||\mu_i| - |\nu_i||} \tau_{|\nu|,|\mu|} \prod_{k=1}^D \chi_d(\mu_k, \nu_k). \tag{5.24}$$

With the choice $\omega_\nu := 2^{-\frac{d}{2} \sum_i |\nu_i|}$, the right hand side of (5.24) can be rewritten as

$$\varepsilon_0\, 2^{\frac{d}{2} \sum_i |\nu_i|} \sum_{j \in \mathbb{Z}^D_{j_0}} \tau_{|\nu|,j} \prod_{k=1}^D 2^{-\frac{d}{2} j_k} \sum_{|\mu_k| = j_k} 2^{-\frac{d}{2} |j_k - |\nu_k||} \chi_d(\mu_k, \nu_k).$$

Using the estimate $\sum_{|\mu_k| = j_k} 2^{-\frac{d}{2} |j_k - |\nu_k||} \chi_d(\mu_k, \nu_k) \lesssim 2^{-\frac{d}{2}(|\nu_k| - j_k)}$, we obtain

$$\omega_\nu^{-1} \sum_{\mu \in (\nabla^{(d)})^D} \omega_\mu |m_{\nu\mu}| s_{|\mu|} s_{|\nu|} \lesssim \varepsilon_0 \sum_{j \in \mathbb{Z}^D_{j_0}} \tau_{|\nu|,j} \leq C_\tau\, \varepsilon_0,$$

which with (5.23) implies the assertion. □

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

Proof of Theorem 5.2. Let $X := \mathrm{H}^s(\mathbb{R}^3)$ and denote by X_Λ the closure of $\operatorname{span}\{\Psi_\nu\}_{\nu\in\Lambda}$ in X. For $u \in X_\Lambda$, in what follows u_ν denotes the corresponding coefficients of the wavelet expansion of u.

The operator norm $\left\| A_\Lambda - \tilde{A}_\Lambda \,\middle|\, \mathrm{H}^s(\mathbb{R}^3) \to \mathrm{H}^{-s}(\mathbb{R}^3) \right\|$ can be estimated by

$$\sup_{\substack{u,v\in X_\Lambda\\ \|u\|_X=\|v\|_X=1}} \langle (A_\Lambda - \tilde{A}_\Lambda)u, v\rangle \le \sup_{\substack{u,v\in X\\ \|u\|_X=\|v\|_X=1}} \int_{B_r(0)} \frac{1}{|x|}\, u\, v \,\mathrm{d}x + \sup_{\substack{u,v\in X_\Lambda\\ \|u\|_X=\|v\|_X=1}} r^{-1}\delta(N_{|\mu|,|\nu|}, r^{-2}R^2) \left(\int_{\mathbb{R}^3\setminus B_r(0)} |\Psi_\nu \Psi_\mu| \,\mathrm{d}x \right) u_\mu\, v_\nu.$$

For the first term on the right hand side, we can apply Proposition 5.3; the estimate for the second term depends on $\bar{\Lambda}$. If $\bar{\Lambda} = \nabla^{(3)}$, using that φ and ψ are uniformly bounded by our regularity assumptions, we have

$$\int_{\mathbb{R}^3} |\Psi_\mu \Psi_\nu| \,\mathrm{d}x \lesssim 2^{-\frac{3}{2}||\mu|-|\nu||} \chi_3(\mu,\nu),$$

and with (5.20), the second term can be estimated by

$$r^{-1}\delta(N_{|\mu|,|\nu|}, r^{-2}R^2) \int_{\mathbb{R}^3\setminus B_r(0)} |\Psi_\nu \Psi_\mu| \,\mathrm{d}x \le \varepsilon_0 2^{s(|\mu|+|\nu|)} \tau_{|\mu|,|\nu|} 2^{-\frac{3}{2}||\mu|-|\nu||} \chi_3(\mu,\nu).$$

Hence Lemma 5.4 gives the assertion. In the case $\bar{\Lambda} = (\nabla^{(1)})^3$, we obtain

$$\int_{\mathbb{R}^3} |\Psi_\mu \Psi_\nu| \,\mathrm{d}x \lesssim 2^{-\frac{1}{2}\sum_i ||\mu_i|-|\nu_i||} \prod_{i=1}^3 \chi_1(\mu_i,\nu_i),$$

and the assertion follows with (5.21) and Lemma 5.4. □

The estimate (5.22) can be transferred to the three-dimensional Coulomb potential acting on higher-dimensional functions, since, *e.g.*, $\mathrm{H}^s(\mathbb{R}^{3+n}) \subset \mathrm{H}^s(\mathbb{R}^3) \otimes \mathrm{L}^2(\mathbb{R}^n)$ for any n by Lemma 5.1.

5.2.4. Approximation of two-electron operators

We now use the same strategy as in the previous subsection for the two-electron Coulomb potential and the modified potential (3.6). For the following, let $S_\rho := \{x, y \in \mathbb{R}^3 :\ |x-y| < \rho\}$ for $\rho > 0$.

Theorem 5.6. *Let $0 < r < R$, $\varepsilon_0 > 0$, and either $s = 1$ or $s \in (\frac{3}{2}, 2]$. For each $N \in \mathbb{N}$, let $p_N := p_{N,r,R}$. Furthermore, let either*

(1) *$\bar{\Lambda} = (\nabla^{(3)})^2$ and $\Lambda \subset \bar{\Lambda}$ such that $\operatorname{supp}\Psi_\nu \subset S_R$ for all $\nu \in \Lambda$; let $\tau \in W_2$ and for $\alpha, \beta \in \mathbb{Z}^2_{j_0}$, let $N_{\alpha,\beta}$ be chosen such that*

$$r^{-1}\delta(N_{\alpha,\beta}, r^{-2}R^2) \le \left(\sum_{k,l=1}^2 2^{2s(\alpha_k+\beta_l)} \right)^{\frac{1}{2}} \tau_{\alpha,\beta}\varepsilon_0;$$

or

(2) *$\bar{\Lambda} = (\nabla^{(1)})^6$ and $\Lambda \subset \bar{\Lambda}$ such that $\operatorname{supp}\Psi_\nu \subset S_R$ for all $\nu \in \Lambda$; let $\tau \in W_6$ and for $\alpha, \beta \in \mathbb{Z}^6_{j_0}$, let $N_{\alpha,\beta}$ be chosen such that*

$$r^{-1}\delta(N_{\alpha,\beta}, r^{-2}R^2) \le \left(\sum_{k,l=1}^6 2^{2s(\alpha_k+\beta_l)} \right)^{\frac{1}{2}} \tau_{\alpha,\beta}\varepsilon_0.$$

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

Let $A_\Lambda, \tilde{A}_\Lambda \colon \mathrm{H}^s(\mathbb{R}^6) \to \mathrm{H}^{-s}(\mathbb{R}^6)$ *be the operators defined by the matrices* $(a_{\nu\mu})_{\nu,\mu\in\Lambda}$, $(\tilde{a}_{\nu\mu})_{\nu,\mu\in\Lambda}$ *with*

$$a_{\nu\mu} = \int_{\mathbb{R}^6} \frac{1}{|x-y|} \Psi_\mu \Psi_\nu \,\mathrm{d}(x,y), \quad \tilde{a}_{\nu\mu} = \int_{\mathbb{R}^6} p_{N_{|\mu|,|\nu|}}(x-y)\, \Psi_\mu \Psi_\nu \,\mathrm{d}(x,y),$$

extended to $\mu,\nu \in \bar{\Lambda}\setminus\Lambda$ *by zero. Then there exist* $C_s, C_{\psi,j_0,s,\tau} > 0$ *independent of* r, R, ε_0 *and* Λ *such that*

$$\left\| A_\Lambda - \tilde{A}_\Lambda \,\middle|\, \mathrm{H}^s(\mathbb{R}^6) \to \mathrm{H}^{-s}(\mathbb{R}^6) \right\| \le C_s\, r^{\eta_s} + C_{\psi,j_0,s,\tau}\, \varepsilon_0, \tag{5.25}$$

where $\eta_1 = 1$, *and* $\eta_s = 2$ *for* $s \in (\frac{3}{2}, 2]$.

Similarly to the previous subsection, the following proposition is used to estimate the error corresponding to the domain S_ρ.

Proposition 5.7. *Let either* $s = 1$ *or* $s > \frac{3}{2}$. *Then there exists* $C > 0$ *such that for* $r > 0$,

$$\sup\left\{ \int_{|x-y|<r} \frac{1}{|x-y|} f\, g \,\mathrm{d}(x,y) :\ f,g \in \mathrm{H}^s(\mathbb{R}^6),\ \|f\|_s = \|g\|_s = 1 \right\} \le C r^{\eta_s},$$

where $\eta_1 = 1$, *and* $\eta_s = 2$ *for* $s > \frac{3}{2}$.

Proof. Follows as in Proposition 5.3 with a coordinate rotation and by $\mathrm{H}^s(\mathbb{R}^6) \subset \mathrm{H}^s(\mathbb{R}^3) \otimes \mathrm{L}^2(\mathbb{R}^3)$ according to Lemma 5.1. □

Proof of Theorem 5.6. Following the lines of the proof of Theorem 5.2, combining Proposition 5.7 with Lemma 5.4. □

For the modified potential (3.6), we additionally obtain from (5.17) the exponential sum approximation

$$\left\| \left| \frac{x}{|x|} - x \sum_{k=1}^{N} \frac{\omega_k}{r} \exp\left(-\frac{\alpha_k}{r^2}|x|^2\right) \right|_{\ell^2} \right\|_{\infty,\{r\le|x|\le R\}} \le r^{-1} R\, \delta(N, r^{-2}R^2). \tag{5.26}$$

Theorem 5.8. *Let* $0 < r < R$, $\varepsilon_0 > 0$, *and either* $s = 1$ *or* $s = 2$. *For each* $N \in \mathbb{N}$, *let* $p_N := p_{N,r,R}$. *Furthermore, let either*

(1) $\bar{\Lambda} = (\nabla^{(3)})^2$ *and* $\Lambda \subset \bar{\Lambda}$ *such that* $\operatorname{supp}\Psi_\nu \subset S_R$ *for all* $\nu \in \Lambda$; *let* $\tau \in W_2$ *and for* $\alpha,\beta \in \mathbb{Z}^2_{j_0}$, *let* $N_{\alpha,\beta}$ *be chosen such that*

$$r^{-1} R\, \delta(N_{\alpha,\beta}, r^{-2}R^2) \le (2^{2\alpha_1} + 2^{2\alpha_2})^{-\frac{1}{2}} \left(\sum_{k,l=1}^{2} 2^{2s(\alpha_k+\beta_l)} \right)^{\frac{1}{2}} \tau_{\alpha,\beta}\varepsilon_0;$$

or

(2) $\bar{\Lambda} = (\nabla^{(1)})^6$ *and* $\Lambda \subset \bar{\Lambda}$ *such that* $\operatorname{supp}\Psi_\nu \subset S_R$ *for all* $\nu \in \Lambda$; *let* $\tau \in W_6$ *and for* $\alpha,\beta \in \mathbb{Z}^6_{j_0}$, *let* $N_{\alpha,\beta}$ *be chosen such that*

$$r^{-1} R\, \delta(N_{\alpha,\beta}, r^{-2}R^2) \le \left(\sum_{k=1}^{6} 2^{2\alpha_k} \right)^{-\frac{1}{2}} \left(\sum_{k,l=1}^{6} 2^{2s(\alpha_k+\beta_l)} \right)^{\frac{1}{2}} \tau_{\alpha,\beta}\varepsilon_0.$$

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

Let $A_\Lambda, \tilde{A}_\Lambda \colon \mathrm{H}^s(\mathbb{R}^6) \to \mathrm{H}^{-s}(\mathbb{R}^6)$ be the operators defined by the matrices $(a_{\nu\mu})_{\nu,\mu\in\Lambda}$, $(\tilde{a}_{\nu\mu})_{\nu,\mu\in\Lambda}$ with

$$a_{\nu\mu} = \int_{\mathbb{R}^6} \frac{x-y}{|x-y|} \cdot (\mathrm{D}_x - \mathrm{D}_y)\, \Psi_\mu\, \Psi_\nu \,\mathrm{d}(x,y),$$

$$\tilde{a}_{\nu\mu} = \int_{\mathbb{R}^6} \left[(x-y)\, p_{N_{|\mu|},|\nu|}(x-y)\right] \cdot (\mathrm{D}_x - \mathrm{D}_y)\, \Psi_\mu\, \Psi_\nu \,\mathrm{d}(x,y),$$

extended to $\mu, \nu \in \bar{\Lambda} \setminus \Lambda$ by zero. Then there exist $C_s, C_{\psi,j_0,s,\tau} > 0$ independent of r, R, ε_0 and Λ such that

$$\left\| A_\Lambda - \tilde{A}_\Lambda \,\middle|\, \mathrm{H}^s(\mathbb{R}^6) \to \mathrm{H}^{-s}(\mathbb{R}^6) \right\| \le C_s\, r^{\eta_s} + C_{\psi,j_0,s,\tau}\, \varepsilon_0, \tag{5.27}$$

where $\eta_1 = 1$, and $\eta_2 = \frac{5}{2}$.

Note the slight difference in powers of r between (5.25) and (5.27), which is related to the following proposition.

Proposition 5.9. *Let either $s = 1$ or $s = 2$, then there exists $C > 0$ such that for $r > 0$,*

$$\sup\left\{ \int_{|x-y|<r} \frac{x-y}{|x-y|} \cdot (\mathrm{D}_x - \mathrm{D}_y) f\, g \,\mathrm{d}(x,y) : \ f, g \in \mathrm{H}^s(\mathbb{R}^6),\ \|f\|_s = \|g\|_s = 1 \right\} \le C r^{\eta_s},$$

where $\eta_1 = 1$, $\eta_2 = \frac{5}{2}$.

Proof. We again use a rotation of coordinates and $\mathrm{H}^s(\mathbb{R}^6) \subset \mathrm{H}^s(\mathbb{R}^3) \otimes \mathrm{L}^2(\mathbb{R}^3)$. In the case $s = 1$, it suffices to estimate, for $f, g \in \mathrm{H}^1(\mathbb{R}^3)$,

$$\int_{B_r(0)} \frac{x}{|x|} \cdot \mathrm{D}_x f\, g \,\mathrm{d}x \le \|f\|_{\mathrm{H}^1(\mathbb{R}^3)} \| |x|^{-1} |x| \|_{\mathrm{L}^3(B_r(0))} \|g\|_{\mathrm{L}^6(\mathbb{R}^3)} \lesssim \left(\int_0^r s^2 \,\mathrm{d}s \right)^{1/3} \|f\|_{\mathrm{H}^1(\mathbb{R}^3)} \|g\|_{\mathrm{H}^1(\mathbb{R}^3)}.$$

In the case $s = 2$, for $f, g \in \mathrm{H}^2(\mathbb{R}^3)$,

$$\int_{B_r(0)} \frac{x}{|x|} \cdot \mathrm{D}_x f\, g \,\mathrm{d}x \lesssim \| |x|^{-1} |x| \|_{\mathrm{L}^{6/5}(B_r(0))} \| |\mathrm{D}_x f| \|_{\mathrm{L}^6(\mathbb{R}^3)} \|g\|_{\mathrm{H}^2(\mathbb{R}^3)} \lesssim \left(\int_0^r s^2 \,\mathrm{d}s \right)^{5/6} \|f\|_{\mathrm{H}^2(\mathbb{R}^3)} \|g\|_{\mathrm{H}^2(\mathbb{R}^3)},$$

where we have used the imbedding $\mathrm{H}^2(\mathbb{R}^3) \hookrightarrow \mathrm{L}^\infty(\mathbb{R}^3)$, Hölder's inequality, and the imbedding $\mathrm{H}^1(\mathbb{R}^3) \hookrightarrow \mathrm{L}^6(\mathbb{R}^3)$. □

Proof of Theorem 5.8. Following the lines of the proof of Theorem 5.2, combining Proposition 5.9 with Lemma 5.4. □

5.2.5. Eigenpair error estimates

We finally come to the error in eigenvalues and eigenfunctions due to the approximation of potentials. The required estimates are first formulated in abstract terms and eventually applied to the concrete approximations we have considered. Note that in the case of the modified problem involving the nonsymmetric term (3.6), although the exact eigenvalues of interest are unchanged by the discussion in Section 2, it cannot necessarily be guaranteed that the eigenvalues remain real for the perturbed potentials, hence we need to work on spaces of complex-valued functions at this point. To avoid technicalities, and since this case is typically found in electronic ground states, we restrict the discussion to simple eigenvalues.

Let $d \in \mathbb{N}$, let $V \subseteq \mathrm{H}^1(\mathbb{R}^d, \mathbb{C})$ be a closed subspace, and let the real linear operators $A, A_n \colon V \to V'$ be bounded and invertible. Let $\lambda_0 \in \mathbb{C}$ be an isolated simple eigenvalue with eigenfunction u_0 of A and let

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

$\|A - A_n\| \le \varepsilon_n \to 0$. In the following, $\langle\cdot,\cdot\rangle$ denotes the duality product induced by the inner product on $\mathrm{L}^2(\mathbb{R}^d,\mathbb{C})$, and $\|\cdot\|$ denotes the norm on $\mathrm{H}^1(\mathbb{R}^d,\mathbb{C})$. We define the operators $T, T_n\colon V \to V$ by

$$\langle ATu, v\rangle = \langle u, v\rangle, \quad \langle A_nT_nu, v\rangle = \langle u, v\rangle \qquad \text{for all } u, v \in V.$$

Then $\mu_0 := \lambda_0^{-1}$ is isolated simple eigenvalue of T, and we can choose a $\delta > 0$ such that $\{|z-\mu_0| \le \delta\} \cap (\sigma(T) \setminus \{\mu_0\}) = \emptyset$; for this δ, let $\Gamma := \{z \in \mathbb{C}\colon |z-\mu_0| = \delta\}$.

Theorem 5.10. *For sufficiently large n, there exist isolated simple eigenvalues $\mu_{0,n}$ of T_n with corresponding eigenfunctions $u_{0,n}$ as well as $C_1, C_2 > 0$ such that*

$$|\mu_0 - \mu_{0,n}| \le C_1\varepsilon_n, \quad \|u_0 - u_{0,n}\| \le C_2\varepsilon_n.$$

Proof. Let $M_\Gamma = \max_{z\in\Gamma}\|(T-z)^{-1}\|$. Without restriction of generality, we can assume $\varepsilon_n M_\Gamma \le \frac{1}{2}$, which implies $\|(T_n - z)^{-1}\| \le 2\|(T-z)^{-1}\| \le 2M_\Gamma$ for all $z \in \Gamma$ and all n, see *e.g.* [35], IV, Theorem 1.16. Thus in particular $\Gamma \subset \rho(T_n)$ for all n. Hence we can define the spectral projections

$$P = -\frac{1}{2\pi i}\int_\Gamma (T-\zeta)^{-1}\,\mathrm{d}\zeta, \quad P_n = -\frac{1}{2\pi i}\int_\Gamma (T_n-\zeta)^{-1}\,\mathrm{d}\zeta$$

and obtain for $u \in \mathrm{H}^1(\mathbb{R}^d,\mathbb{C})$, following the lines of [9], Proposition 5.3,

$$\|(P-P_n)u\| = \frac{1}{2\pi}\left\|\int_\Gamma (T_n-\zeta)^{-1}(T_n-T)(T-\zeta)^{-1}u\,\mathrm{d}\zeta\right\| \le 2\delta M_\Gamma^2\,\varepsilon_n\|u\|,$$

and hence $\|P - P_n\| \to 0$.

From $\|P-P_n\| < 1$ it follows that $\dim\operatorname{im} P = \dim\operatorname{im}\tilde{P}_n = 1$, see [9], page 87. Consequently, by [35], page 182, for large enough n, T_n has an isolated simple eigenvalue $\mu_{0,n}$ inside Γ with corresponding eigenfunction $u_{0,n}$, normalized as u_0, such that for an n-independent $C_2 > 0$, $\|u_0 - u_{0,n}\| \le C_2\varepsilon_n$. The identity

$$T(u_0 - u_{0,n}) + (T-T_n)u_{0,n} = \mu_0(u_0 - u_{0,n}) + (\mu_0 - \mu_{0,n})u_{0,n}$$

now yields $|\mu_0 - \mu_{0,n}| \le C_1\varepsilon_n$ for a $C_1 > 0$. □

Remark 5.11. Theorem 5.10 applies also to the adjoints T^*, T_n^*, *i.e.*, if u_0^* is the adjoint eigenfunction for $\mu_0 = \lambda_0^{-1}$, we obtain $u_{0,n}^*$ with $\|u_0^* - u_{0,n}^*\| \le C_2\varepsilon_n$ as well.

Note furthermore that if $\varepsilon_n \le (2C_1|\lambda_0|)^{-1}$, the estimate $|\mu_0 - \mu_{0,n}| \le C_1\varepsilon_n$ implies $|\lambda_{0,n}| \le 2|\lambda_0|$ and hence $|\lambda_0 - \lambda_{0,n}| \le 2C_1|\lambda_0|^2\varepsilon_n$.

Proposition 5.12. *Let $A, \tilde{A}\colon \mathrm{H}^1(\mathbb{R}^d,\mathbb{C}) \to \mathrm{H}^{-1}(\mathbb{R}^d,\mathbb{C})$ be bounded with*

$$\sup\left\{\langle (A-\tilde{A})u, v\rangle :\ \|u\|_{\mathrm{H}^2(\mathbb{R}^d,\mathbb{C})} = \|v\|_{\mathrm{H}^2(\mathbb{R}^d,\mathbb{C})} = 1\right\} \le \varepsilon.$$

Moreover, let $u, \tilde{u}_0, u_0^*, \tilde{u}_0^* \in \mathrm{H}^2(\mathbb{R}^d,\mathbb{C})$ and $\lambda_0, \tilde{\lambda}_0 \in \mathbb{C}$ with $Au_0 = \lambda_0 u_0$, $\tilde{A}\tilde{u}_0 = \tilde{\lambda}_0\tilde{u}_0$, $A^*u_0^* = \overline{\lambda_0}u_0^*$, $\tilde{A}^*\tilde{u}_0^* = \overline{\tilde{\lambda}_0}\tilde{u}_0^*$, with the normalizations $\langle u_0, u_0^*\rangle = \langle\tilde{u}_0, \tilde{u}_0^*\rangle = 1$. Then

$$|\lambda_0 - \tilde{\lambda}_0| \lesssim \|u_0 - \tilde{u}_0\|\,\|u_0^* - \tilde{u}_0^*\| + \varepsilon\,\|\tilde{u}_0\|_{\mathrm{H}^2(\mathbb{R}^d,\mathbb{C})}\,\|\tilde{u}_0^*\|_{\mathrm{H}^2(\mathbb{R}^d,\mathbb{C})}. \tag{5.28}$$

Proof. We have

$$\begin{aligned}\langle A(u_0-\tilde{u}_0), u_0^* - \tilde{u}_0^*\rangle &= \langle Au_0, u_0^*\rangle + \langle A\tilde{u}_0, \tilde{u}_0^*\rangle - \langle Au_0, \tilde{u}_0^*\rangle - \langle A\tilde{u}_0, u_0^*\rangle\\ &= \lambda_0 + \tilde{\lambda}_0 + \langle (A-\tilde{A})\tilde{u}_0, \tilde{u}_0^*\rangle - \lambda_0(\langle u_0, \tilde{u}_0^*\rangle + \langle\tilde{u}_0, u_0^*\rangle)\\ &= -\lambda_0 + \tilde{\lambda}_0 + \langle (A-\tilde{A})\tilde{u}_0, \tilde{u}_0^*\rangle + \lambda_0\langle u_0 - \tilde{u}_0, u_0^* - \tilde{u}_0^*\rangle,\end{aligned}$$

and therefore $|\lambda_0 - \tilde{\lambda}_0| \le \|A - \lambda_0\,\mathrm{I}\|\,\|u_0 - \tilde{u}_0\|\,\|u_0^* - \tilde{u}_0^*\| + \varepsilon\,\|\tilde{u}_0\|_{\mathrm{H}^2(\mathbb{R}^d,\mathbb{C})}\,\|\tilde{u}_0^*\|_{\mathrm{H}^2(\mathbb{R}^d,\mathbb{C})}$. □

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

Remark 5.13. For self-adjoint operators, it suffices to consider real-valued function spaces and real eigenvalues. With the additional assumption that λ_0, $\tilde{\lambda}_0$ are the lowest eigenvalues, and if $u_0, \tilde{u}_0 \in \mathrm{H}^2$, the characterization by the Rayleigh quotient yields $|\lambda_0 - \tilde{\lambda}_0| \leq \varepsilon \max\{\|u_0\|^2_{\mathrm{H}^2(\mathbb{R}^d)}, \|\tilde{u}_0\|^2_{\mathrm{H}^2(\mathbb{R}^d)}\}$.

We now apply Theorem 5.10 to the approximations from Theorems 5.2, 5.6 and 5.8, where a potential operator A_Λ is approximated by a family of separable substitutes $\tilde{A}_\Lambda$ with error in operator norm dependent on the two parameters r, ε_0, the latter being tied to the rank of the separable expansions by (5.18), (5.26). Thus the Hamiltonians formed with the respective A_Λ need to be invertible, which can be ensured by adding appropriate shifts, and the eigenvalues of interest need to be isolated and simple; then Theorem 5.10 yields

$$|\lambda_0 - \tilde{\lambda}_0| \lesssim r + \varepsilon_0, \quad \|u_0 - \tilde{u}_0\|_{\mathrm{H}^1} \lesssim r + \varepsilon_0, \quad \|u_0^* - \tilde{u}_0^*\|_{\mathrm{H}^1} \lesssim r + \varepsilon_0,$$

where u_0, u_0^* can be direct and adjoint eigenfunctions of discretizations of (3.2) or of (3.3) according to the considered A_Λ, and $\tilde{u}_0$, $\tilde{u}_0^*$ those obtained with approximate potentials. An estimate of the same order in r, ε_0 for $|\lambda_0 - \tilde{\lambda}_0|$ that applies to (3.2) has also been given in [51].

At this point we can make use of the assumption made in the beginning of Section 4 that the underlying univariate wavelet basis is H^2-regular: since if we know that $u_0, u_0^* \in \mathrm{H}^2$, which is the case by regularity of the exact solutions and $\varphi, \psi \in \mathrm{H}^2$, and that $\|\tilde{u}_0\|_{\mathrm{H}^2}, \|\tilde{u}_0^*\|_{\mathrm{H}^2} \leq C$ with some $C > 0$ independent of the approximation parameters, which can be inferred from $\varphi, \psi \in \mathrm{H}^2$ and operator norm convergence of the separable approximation, then (5.28) and Theorems 5.2, 5.6, and 5.8 additionally yield

$$|\lambda_0 - \tilde{\lambda}_0| \lesssim (r + \varepsilon_0)^2 + r^2 + \varepsilon_0.$$

In the symmetric case, if λ_0, $\tilde{\lambda}_0$ are the lowest eigenvalues, by Remark 5.13, we have $|\lambda_0 - \tilde{\lambda}_0| \lesssim r^2 + \varepsilon_0$.

In other words, if the basis functions are smooth enough and if one chooses $\varepsilon_0 \sim r^2$, the error in energy induced by the separable approximations depends not linearly, but quadratically on r. In view of the estimates (5.18), (5.26), this yields a substantial improvement in the ranks of the separable approximations required for a certain error in energy.

6. Conclusion

The mixed regularity estimates we have discussed clearly show the improvement in terms of convergence rates of explicitly correlated methods over a direct discretization by hyperbolic wavelets. A second point that is crucial for the practical feasibility of wavelet discretizations is the exploitation of separability of functions where possible, which in particular leads to the separable approximations of potentials that we have discussed in detail. In the error analysis for these approximations, wavelet bases that are at least H^2-regular have turned out to be advantageous for the convergence of the energy.

Concerning the construction of the tensor product wavelet bases, we have noted that working with an orthonormal basis is essentially a requirement. Furthermore we have compared two relevant types of tensor product wavelet constructions and found that tensor products of univariate wavelet bases are preferable to tensor products of isotropically supported three-dimensional basis functions from an algorithmic point of view.

Various different strategies are possible for employing the concepts discussed here in eigenvalue solvers, where adaptive methods will be important for practical implementations. It should be noted that due to nonsymmetry of the transcorrelated problem and lack of compact imbeddings of the involved function spaces, even the convergence of Galerkin discretizations with wavelet indices fixed a priori is not covered by standard results. In a separate work we will consider solvers that build on the techniques that we have studied.

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.

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

Acknowledgements. Financial support from Deutsche Forschungsgemeinschaft through grant GSC 111 is gratefully acknowledged. The author would like to thank Wolfgang Dahmen, Boris Khoromskij, Joachim Schöberl, Karen Veroy, and Andreas Zeiser for helpful discussions and hints.

References

[1] R.A. Adams and J.J.F. Fournier, Sobolev Spaces, *Pure and Applied Mathematics Series*, 2nd edition. Academic Press **140** (2003).

[2] S. Agmon, 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).

[3] M. Bachmayr, *Integration of products of Gaussians and wavelets with applications to electronic structure calculations*. Preprint AICES, RWTH Aachen (2012).

[4] R. Balder and C. Zenger, The solution of multidimensional real Helmholtz equations on sparse grids. *SIAM J. Sci. Comput.* **17** (1996) 631–646.

[5] G. Beylkin, On the representation of operators in bases of compactly supported wavelets. *SIAM J. Numer. Anal.* **29** (1992) 1716–1740.

[6] S.F. Boys and N.C. Handy, The determination of energies and wavefunctions with full electronic correlation. *Proc. R. Soc. Lond. A* **310** (1969) 43–61.

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

[8] H.-J. Bungartz, *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung*. Ph.D. thesis, Technische Universität München (1992).

[9] F. Chatelin, Spectral Approximation of Linear Operators, *Computer Science and Applied Mathematics*. Academic Press (1983).

[10] S.R. Chinnamsetty, M. Espig, B.N. Khoromskij, W. Hackbusch and H.-J. Flad, Tensor product approximation with optimal rank in quantum chemistry. *J. Chem. Phys.* **127** (2007) 084110.

[11] A. Cohen, Numerical Analysis of Wavelet Methods. *Stud. Math. Appl.* **32** (2003).

[12] W. Dahmen and C.A. Micchelli, Using the refinement equation for evaluating integrals of wavelets. *SIAM J. Numer. Anal.* **30** (1993) 507–537.

[13] I. Daubechies, Orthonormal bases of compactly supported wavelets. *Commun. Pure Appl. Math.* **41** (1988) 909–996.

[14] T.J. Dijkema, C. Schwab and R. Stevenson, An adaptive wavelet method for solving high-dimensional elliptic PDEs. *Constr. Approx.* **30** (2009) 423–455.

[15] G. Donovan, J. Geronimo and D. Hardin, Intertwining multiresolution analyses and the construction of piecewise-polynomial wavelets. *SIAM J. Math. Anal.* **27** (1996) 1791–1815.

[16] G. Donovan, J. Geronimo and D. Hardin, Orthogonal polynomials and the construction of piecewise polynomial smooth wavelets. *SIAM J. Math. Anal.* **30** (1999) 1029–1056.

[17] H.-J. Flad, W. Hackbusch, D. Kolb and R. Schneider, Wavelet approximation of correlated wave functions. I. Basics. *J. Chem. Phys.* **116** (2002) 9641–9657.

[18] H.-J. Flad, W. Hackbusch and R. Schneider, Best N-term approximation in electronic structure calculations. II. Jastrow factors. *ESAIM: M2AN* **41** (2007) 261.

[19] H.-J. Flad, W. Hackbusch, B.N. Khoromskij and R. Schneider, Matrix Methods: Theory, Algorithms and Applications, in *Concepts of Data-Sparse Tensor-Product Approximation in Many-Particle Modelling*. World Scientific (2010) 313–347.

[20] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Østergaard Sørensen, Sharp regularity results for many-electron wave functions. *Commun. Math. Phys.* **255** (2005) 183–227.

[21] L. Genovese, T. Deutsch, A. Neelov, S. Goedecker and G. Beylkin, Efficient solution of poisson's equation with free boundary conditions. *J. Chem. Phys.* **125** (2006) 074105.

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

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

[24] M. Griebel and J. Hamaekers, A wavelet based sparse grid method for the electronic Schrödinger equation, in *Proc. of the International Congress of Mathematicians*, edited by M. Sanz-Solé, J. Soria, J. Varona and J. Verdera **III** (2006) 1473–1506.

[25] M. Griebel and J. Hamaekers, Tensor product multiscale many-particle spaces with finite-order weights for the electronic Schrödinger equation. *Z. Phys. Chem.* **224** (2010) 527–543. Also available as INS Preprint No. 0911.

[26] M. Griebel and S. Knapek, Optimized tensor-product approximation spaces. *Constr. Approx.* **16** (2000) 525.

[27] M. Griebel and P. Oswald, Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems. *Adv. Comput. Math.* **4** (1995) 171–206.

[28] J. Hamaekers, *Tensor Product Multiscale Many-Particle Spaces with Finite-Order Weights for the Electronic Schödinger Equation*. Ph.D. thesis, Universität Bonn (2009).

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press

[29] H. Harbrecht, R. Schneider and C. Schwab, Multilevel frames for sparse tensor product spaces. *Numer. Math.* **110** (2008) 199–220.
[30] R.J. Harrison, G.I. Fann, T. Yanai, Z. Gan and G. Beylkin, Multiresolution quantum chemistry: basic theory and initial applications. *J. Chem. Phys.* **121** (2004) 11587–11598.
[31] E. Hille, A class of reciprocal functions. *Ann. Math.* **27** (1926) 427–464.
[32] J.O. Hirschfelder, Removal of electron-electron poles from many-electron Hamiltonians. *J. Chem. Phys.* **39** (1963) 3145–3146.
[33] E. Hylleraas, Über den Grundzustand des Heliumatoms. *Z. Phys.* **48** (1929) 469.
[34] T. Kato, On the eigenfunctions of many-particle systems in quantum mechanics. *Commun. Pure Appl. Math.* **X** (1957) 151–177.
[35] T. Kato, Perturbation Theory for Linear Operators, *Grundlehren der mathematischen Wissenschaften*, 2nd edition. Springer-Verlag, Berlin, Heidelberg, New York **132** (1976).
[36] W. Klopper, R12 methods, Gaussian geminals, in *Modern Methods and Algorithms of Quantum Chemistry*, edited by J. Grotendorst (2000) 181–229.
[37] H.-C. Kreusler and H. Yserentant, *The mixed regularity of electronic wave functions in fractional order and weighted Sobolev spaces.* Preprint 94, DFG SPP 1324 (2011).
[38] H. Luo, D. Kolb, H.-J. Flad, W. Hackbusch and T. Koprucki, Wavelet approximation of correlated wave functions. II. Hyperbolic wavelets and adaptive approximation schemes. *J. Chem. Phys.* **117** (2002) 3625–3638.
[39] Y. Meyer and R. Coifman, Wavelets: Calderon-Zygmund and multilinear operators, Cambridge Studies in *Advanced Mathematics*. Cambridge University Press (1997).
[40] A. Neelov and S. Goedecker, An efficient numerical quadrature for the calculation of the potential energy of wavefunctions expressed in the Daubechies wavelet basis. *J. Comput. Phys.* **217** (2006) 312–339.
[41] M. Nooijen and R.J. Bartlett, Elimination of Coulombic infinities through transformation of the Hamiltonian. *J. Chem. Phys.* **109** (1998).
[42] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Analysis of Operators* IV. Academic Press (1978).
[43] C. Schwab and R.A. Todor, Sparse finite elements for stochastic elliptic problems – higher order moments. *Computing* **71** (2003) 43–63.
[44] R. Stevenson, On the compressibility of operators in wavelet coordinates. *SIAM J. Math. Anal.* **35** (2004) 1110–1132.
[45] W. Sweldens, The lifting scheme: a custom-design construction of biorthogonal wavelets. *Appl. Comput. Harmon. Anal.* **3** (1996) 186–200.
[46] S. Tenno, A feasible transcorrelated method for treating electronic cusps using a frozen Gaussian geminal. *Chem. Phys. Lett.* **330** (2000) 169–174.
[47] D.P. Tew and W. Klopper, New correlation factors for explicitly correlated electronic wave functions. *J. Chem. Phys.* **123** (2005) 074101.
[48] H. Yserentant, On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives. *Numer. Math.* **98** (2004) 731–759.
[49] H. Yserentant, Regularity and Approximability of Electronic Wave Functions. *Lect. Notes Math.* **2000** (2010).
[50] H. Yserentant, The mixed regularity of electronic wave functions multiplied by explicit correlation factors. *ESAIM: M2AN* **45** (2011) 803–824.
[51] A. Zeiser, *Direkte Diskretisierung der Schrödingergleichung auf dünnen Gittern*. Ph.D. thesis, TU Berlin (2010).
[52] A. Zeiser, Fast matrix-vector multiplication in the sparse-grid Galerkin method. *J. Sci. Comput.* **47** (2010) 328–346.
[53] A. Zeiser, Wavelet approximation in weighted Sobolev spaces of mixed order with applications to the electronic Schrödinger equation. To appear in *Constr. Approx.* (2011) DOI: 10.1007/s00365-011-9138-7.
[54] H.J.A. Zweistra, C.C.M. Samson and W. Klopper, Similarity-transformed Hamiltonians by means of Gaussian-damped inter-electronic distances. *Collect. Czech. Chem. Commun.* **68** (2003) 374–386.

https://doi.org/10.1051/m2an/2012009 Published online by Cambridge University Press