Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Best N-term approximation in electronic structure calculations. II. Jastrow factors

by

Heinz-Jürgen Flad, Wolfgang Hackbusch, and Reinhold Schneider

Best N-term approximation in electronic structure calculations. II. Jastrow factors

Heinz-Jürgen Flad and Wolfgang Hackbusch

Max-Planck-Institut für Mathematik in den Naturwissenschaften,

Inselstr. 22-26, D-04103 Leipzig, Germany,

Reinhold Schneider

Christian-Albrechts-Universität Kiel,

Christian-Albrechts-Platz 4, D-24098 Kiel, Germany

Abstract

We present a novel application of best N-term approximation theory in the framework of electronic structure calculations. The paper focus on the description of electron correlations within a Jastrow-type ansatz for the wavefunction. As a starting point we discuss certain natural assumptions on the asymptotic behaviour of two-particle correlation functions $\mathcal{F}^{(2)}$ near electron-electron and electron-nuclear cusps. Based on Nitsche's characterization of best N-term approximation spaces $A_q^\alpha(H^1)$, we prove that $\mathcal{F}^{(2)} \in A_q^\alpha(H^1)$ for $q > 1$ and $\alpha = \frac{1}{q} - \frac{1}{2}$ with respect to a certain class of anisotropic wavelet tensor product bases. Computational arguments are given in favour of this specific class compared to other possible tensor product bases. Finally, we compare the approximation properties of wavelet bases with standard Gaussian-type basis sets frequently used in quantum chemistry.

Keywords. Best N-term approximation, wavelets, electron correlations, Jastrow factor
Mathematics Subject Classification. 41A50, 41A63, 65Z05, 81V70

1 Introduction

A natural starting point in quantum many-particle theory is to choose an approximate ansatz for the exact wavefunction which is convenient from a technical point of view and incorporates the essential physical features of the problem under consideration. In general it is not possible to make rigorous statements concerning the systematic error related to a specific ansatz for the wavefunction. In a second step an appropriate discretization scheme has to be chosen that enables a numerical solution of the problem. The discretization error is usually better accessible to analysis and provides important insights concerning the computational complexity of many-particle methods. Within the present work, we focus on the Jastrow ansatz for electronic wavefunctions

$$\Psi\left(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\right) = \mathcal{F}\left(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\right) \Phi\left(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\right), \tag{1.1}$$

where the wavefunction consists of a product of a correlation factor $\mathcal{F}$, usually called *Jastrow factor*, and a given approximate solution Φ. In most applications, Φ corresponds to a single Slater determinant, obtained, e.g., from a Hartree-Fock calculation which incorporates the permutational symmetry of the wavefunction. Therefore, Jastrow factors are required to be symmetric with respect to an interchange of electron coordinates $\mathbf{x}_i \in \mathbb{R}^3$. It is convenient to express the Jastrow factor in exponential form

$$\mathcal{F}\left(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\right) = \exp\left(\sum_i \mathcal{F}^{(1)}(\mathbf{x}_i) + \sum_{i<j} \mathcal{F}^{(2)}(\mathbf{x}_i, \mathbf{x}_j) + \cdots\right), \tag{1.2}$$

where it decomposes into p-particle correlation functions $\mathcal{F}^{(p)}$ with $p = 1, 2, \ldots, n$. These functions possess the cluster property (cf. [4]), which means that for an arbitrary subset of p particles $P := \{i_1, ..., i_p\}$, the

correlation function vanishes if at least one of the particles is separated from the others, i.e.,

$$\mathcal{F}^{(p)}(\mathbf{x}_{i_1}, ..., \mathbf{x}_{i_p}) \rightarrow 0 \qquad \text{as} \quad \max\{|\mathbf{x}_{i_a} - \mathbf{x}_{i_b}| : i_a, i_b \in P\} \rightarrow \infty.$$

For most applications it is possible to neglect correlation functions beyond a certain order (cf. [4]). With this, the exponential ansatz guarantees size-consistency for extended systems[1] which means that expectation values of operators that correspond to extensive thermodynamic properties, like the energy, are proportional to the system size.

Within the present work we consider the nonrelativistic Hamiltonian in atomic units,

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

where $\mathbf{x}_i \in \mathbb{R}^3$ are the electron coordinates and Z_a, $\mathbf{R}_a$ are charges and positions of the nuclei, respectively. The corresponding expectation value of the energy for the product ansatz (1.1) is given by

$$E = \frac{\langle \mathcal{F}\Phi, H\mathcal{F}\Phi \rangle}{\langle \mathcal{F}\Phi, \mathcal{F}\Phi \rangle}, \tag{1.3}$$

where $\langle \; , \; \rangle$ denotes the scalar product in the $L^2(\mathbb{R}^{3n})$ Hilbert space of many-particle wavefunctions. A variational theory of Jastrow factors requires the minimization of the expectation value of the energy (1.3) with respect to the correlation functions $\mathcal{F}^{(1)}, \mathcal{F}^{(2)}, \ldots$. Even for the simplest systems this is a formidable task due to the nonlinear character of the Jastrow ansatz and the high dimensionality of the resulting integrals. A direct numerical treatment is therefore only possible in *quantum Monte Carlo* (QMC) methods where efficient algorithms have been developed for a stochastic optimization of Jastrow factors (cf. [42, 48, 49]).

In order to develop traditional many-particle theories for the Jastrow ansatz, which result in computationally tractable problems, it turns out to be necessary to introduce further approximations. For most applications, including electronic structure calculations, a two-particle correlation function $\mathcal{F}^{(2)}$ of Jastrow factors provides the dominant contribution. Possible contributions from one-particle terms $\mathcal{F}^{(1)}$ can be incorporated into the Slater determinant of the product ansatz (1.1). The remaining contributions from three-particle and higher order terms are fairly small (cf. [30]) because of Pauli's principle which forces the wavefunction to vanish at the coalescence points of three or more electrons. Therefore, most of the developments in many-particle theory restrict to two-particle correlation functions $\mathcal{F}^{(2)}$. These theories, we just want to mention the *Fermi hypernetted chain* (FHNC) method (cf. [4]), have reached a high level of sophistication for homogeneous systems in condensed matter and nuclear physics. A generalization of the FHNC method to inhomogeneous systems has been developed by Krotscheck et al. [2, 33, 34, 41]. The resulting integro-differential equations are still difficult to handle in numerical calculations, which presently limits its applicability to systems with additional symmetries like slabs or films that are homogeneous in two dimensions (cf. [33]) or spherically symmetric atoms (cf. [2, 41]). Nevertheless, the FHNC method can achieve an overall complexity similar to *density functional theory* (DFT). The basic quantities which have to be considered are two-particle functions related to the two-particle correlation function $\mathcal{F}^{(2)}$ and the reduced one-particle density matrix of the reference wavefunction Φ. This has to be compared with the Kohn-Sham formalism in DFT, where the density matrix alone is already sufficient. If it is possible to achieve an efficient numerical treatment for these two-particle functions, both methods are, at least from a formal point of view, of compatible complexity. The advantage of the FHNC method is the underlying many-body theory which is based on a systematic hierarchy of approximations.

For a large range of applications it is sufficient to consider linearized versions of the original variational problem. Such kind of approaches comprise the local ansatz of Stollhoff-Fulde [19, 44, 45], Talman's [46, 47] cluster expansion and a recently proposed Jastrow-Møller-Plesset perturbation theory (cf. [38]). These methods can be expressed in terms of standard Galerkin schemes for the two-particle correlation function $\mathcal{F}^{(2)}$. Given a basis $\Sigma_\Lambda := \{\chi_\lambda : \lambda \in \Lambda\}$ for a finite index set Λ, the Galerkin scheme for the local ansatz corresponds to a $\mathcal{F}_\Lambda^{(2)} \in \text{span}\{\Sigma_\Lambda\}$ that is a solution of the equation

$$\langle \chi_\lambda \Phi, (H - H_{\text{SCF}})\Phi \rangle + \langle \chi_\lambda \Phi, H\mathcal{F}_\Lambda^{(2)}\Phi \rangle - \langle \chi_\lambda \Phi, \mathcal{F}_\Lambda^{(2)}\Phi \rangle \langle \Phi, H\Phi \rangle = 0 \qquad \text{for all } \chi_\lambda \in \Sigma_\Lambda.$$

---

[1]In the terminology of quantum chemistry this is usually called size-extensivity.

For simplicity, the local ansatz has been defined with respect to the Hartree-Fock model (cf. [19]), where the Hartree-Fock Hamiltonian H_{SCF} and Slater determinant Φ enters into the Galerkin scheme. The correlation energy can be calculated from the Galerkin solution via

$$E_{\text{corr}} = \langle \Phi, (H - H_{\text{SCF}}) \mathcal{F}_\Lambda^{(2)} \Phi \rangle.$$

It is obvious from the high dimensional character of the problem that an efficient discretization requires a careful choice of the basis set. Within the present work, we focus on wavelet bases which seem to be a natural candidate from the physical point of view (cf. [12]). An illustrative example for this can be found in Appendix A, where we present a multi-resolution analysis for a Jastrow factor in the case of a homogeneous electron gas. In previous papers [12, 13, 37] we have discussed an "optimal" construction of wavelet bases based on empirical evidence. It is the purpose of the present work to present more rigorous arguments based on certain natural assumptions concerning the regularity of Jastrow factors. The computational complexity does not only depend on the cardinality of the basis set but also on its specific properties which enter into the computation of the Galerkin matrix. This topic is beyond the scope of our paper and we refer to [13] for a detailed discussion.

1.1 Basic assumptions on the regularity of Jastrow factors

In order to study the approximation of the two-particle correlation functions $\mathcal{F}^{(2)}$ in a wavelet bases, we have to assume certain smoothness properties. Proceeding on the assumption that $\mathcal{F}^{(2)}$ is smooth away from the diagonal, with possible exceptions at the positions of the nuclei, we have to specify the asymptotic behaviour of $\mathcal{F}^{(2)}$ and its derivatives near the *electron-electron* $(e-e)$ and *electron-nuclear* $(e-n)$ cusps. Both from a physical and mathematical point of view, it makes sense to consider first the case of a pure $e-e$ cusp. This is rigorously justified only for many-particle models where singular nuclear potentials are absent, like jellium models or soft-core pseudo-potentials that are frequently used in solid state physics. For such kind of models, we make the following assumption on the asymptotic smoothness near the $e-e$ cusp.

Assumption 1.1 *The two-particle correlation function $\mathcal{F}^{(2)}$ belongs to $C^\infty(\mathbb{R}^3 \times \mathbb{R}^3 \setminus D)$ with $D := \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 : \mathbf{x} = \mathbf{y}\}$. Furthermore it satisfies the asymptotic smoothness property*

$$|\partial_{\mathbf{x}}^\alpha \partial_{\mathbf{y}}^\beta \mathcal{F}^{(2)}(\mathbf{x}, \mathbf{y})| \le c_{\alpha,\beta} \, |\mathbf{x} - \mathbf{y}|^{1-|\alpha|-|\beta|} \qquad \textit{for } \mathbf{x} \ne \mathbf{y} \textit{ and } |\alpha| + |\beta| \ge 1, \tag{1.4}$$

in any bounded domain $\Omega \times \Omega \subset \mathbb{R}^3 \times \mathbb{R}^3$.

Here we have introduced the usual short-hand notation

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

with absolute value of the multi-index $|\beta| := \beta_1 + \beta_2 + \beta_3$. Due to the complicated character of the variational problem and its approximate variants, it is not possible to give a rigorous justification for our basic assumption even in the case of a homogeneous electron gas. We refer however to Appendix A, where the asymptotic smoothness property is demonstrated for a homogeneous electron gas within the random-phase approximation using a simple argument related to the calculus of pseudo-differential operators. The basic Assumption 1.1 is consistent with the asymptotic analysis of the exact wavefunction near the $e-e$ cusp (cf. [31, 28, 29]). Furthermore there exists a huge amount of empirical evidence from QMC calculations. The standard ansatz in QMC for pure two-electron correlation functions consists of polynomials (cf. [42]) or Padé approximants (cf. [48]) in the inter-electron distance, i.e.,

$$\mathcal{F}^{(2)}(\mathbf{x}_i, \mathbf{x}_j) = \sum_\ell d_\ell \, r_{ij}^\ell \quad \text{or} \quad \mathcal{F}^{(2)}(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_\ell b_\ell \, r_{ij}^\ell}{1 + \sum_\ell d_\ell \, r_{ij}^\ell} \qquad \text{with } r_{ij} := |\mathbf{x}_i - \mathbf{x}_j|,$$

respectively, which obviously satisfy the basic Assumption 1.1. In order to avoid a blow up of the polynomial ansatz at $r_{ij} \to \infty$, either a renormalized inter-electron distance $\overline{r}_{ij}$ with $\lim_{r_{ij} \to \infty} \overline{r}_{ij} \to 1$ is used (cf. [42]) or the polynomial is multiplied by an appropriate cut-off function (cf. [49]). For inhomogeneous systems,

like atoms and molecules, a significant amount of the correlation energy is due to *electron-electron-nuclear* $(e-e-n)$ correlations, which are described by incorporating electron-nuclear distances into the polynomials (cf. [42, 48]), e.g.,

$$\mathcal{F}^{(2)}(\mathbf{x}_i, \mathbf{x}_j) = \sum_\alpha \sum_\ell d^\alpha_{\ell mn} r^\ell_{ij} s^m_{\alpha,i} s^n_{\alpha,j} \qquad \text{with } s_{\alpha,i} := |\mathbf{x}_i - \mathbf{R}_\alpha|, \tag{1.5}$$

where $\mathbf{R}_\alpha$ corresponds to the positions of the nuclei. This ansatz was strongly motivated by the highly accurate Hylleraas wavefunctions for two-electron systems (cf. [18]). Such kind of polynomial expansion raises the question for off-diagonal $e-n$ cusps and the behaviour of $\mathcal{F}^{(2)}$ at the three-particle $e-e-n$ cusp. There exist two different kind of approaches to this problem. In the first kind of approach it is required that the approximate solution Φ of the product ansatz (1.1) already satisfies Kato's [31] $e-n$ cusp condition, which means that the leading order term of an asymptotic expansion of the spherically averaged exact wavefunction is correctly reproduced. Neglecting the anisotropy of the wavefunction around a nucleus and higher order terms, it is assumed that $\mathcal{F}^{(2)}$ has no off-diagonal $e-n$ cusps (cf. [42, 16]). As a consequence for coupling terms like $r^\ell_{ij} s^m_{\alpha,i}$ ($\ell, m > 0$) only even powers were employed in the extended ansatz (1.5). No such restrictions are imposed in the second kind of approach (cf. [48]), where all possible combinations are taken into account. The recent comparative study [11] reveals that the second type of approach is only slightly more accurate with respect to the energy, however, provides a considerable improvement in the variance of the local energy which measures the deviation of an approximate wavefunction from an eigenfunction of the Hamiltonian. Therefore it is desirable to consider an extended version of our basic assumption which allows for off-diagonal $e-n$ cusps as well.

Assumption 1.2 *The two-particle correlation function $\mathcal{F}^{(2)}$ belongs to $C^\infty(\mathbb{R}^3 \times \mathbb{R}^3 \setminus (D \cup N))$ with $D := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 : \mathbf{x} = \mathbf{y} \right\}$ and $N := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 : \mathbf{x} = \mathbf{R} \vee \mathbf{y} = \mathbf{R} \right\}$, where $\mathbf{R}$ corresponds to the position of a nucleus. Furthermore it satisfies the generalized asymptotic smoothness property*

$$|\partial^\alpha_\mathbf{x} \partial^\beta_\mathbf{y} \mathcal{F}^{(2)}(\mathbf{x}, \mathbf{y})| \leq c_{\alpha,\beta} \sup_{\alpha_1, \beta_1} \left(\delta_{|\alpha_1|} + |\mathbf{x} - \mathbf{R}|^{1-|\alpha_1|} \right) \left(\delta_{|\beta_1|} + |\mathbf{y} - \mathbf{R}|^{1-|\beta_1|} \right)$$

$$\times \left(\delta_{|\alpha_2|+|\beta_2|} + |\mathbf{x} - \mathbf{y}|^{1-|\alpha_2|-|\beta_2|} \right) \qquad \text{with } \alpha_1 + \alpha_2 = \alpha,\ \beta_1 + \beta_2 = \beta,$$

for $(\mathbf{x}, \mathbf{y}) \notin D \cup N$, in any bounded neighbourhood $\Omega \times \Omega \subset \mathbb{R}^3 \times \mathbb{R}^3$ of the nuclear position $\mathbf{R}$.

This assumption seems to be general enough even for the $e-e-n$ cusp where additional logarithmic terms have to be taken into account (cf. [17, 18]).

So far our discussion focused exclusively on two-particle correlation functions $\mathcal{F}^{(2)}$. However for actual calculations within many-particle methods it might be of interest to consider more general functions including pointwise products with density matrices $\rho(\mathbf{x}, \mathbf{y}) \mathcal{F}^{(2)}(\mathbf{x}, \mathbf{y})$, $\rho(\mathbf{x}, \mathbf{y})^2 \mathcal{F}^{(2)}(\mathbf{x}, \mathbf{y})$ (cf. [13]), or nonlinear expressions like

$$h(\mathbf{x}, \mathbf{y}) = e^{2\mathcal{F}^{(2)}(\mathbf{x}, \mathbf{y})} - 1.$$

The latter replaces $\mathcal{F}^{(2)}$ as the basic quantity in FHNC methods or certain types of cluster expansions (cf. [4]). It is evident that provided $\mathcal{F}^{(2)}$ satisfies Assumptions 1.1 or 1.2 and ρ is the corresponding assumption from [14], the same is true for these functions. Our discussion of best N-term approximations for $\mathcal{F}^{(2)}$ is therefore immediately transferable to such cases.

2 Best N-term approximation for Jastrow factors

Within the last decade, the rather abstract concept of best N-term approximation attracted renewed interest in the context of wavelet based multi-resolution analysis. Initiated by the work of DeVore-Jawerth-Popov [9] a powerful theory has been developed which is based on certain peculiar features of wavelet bases. In order to make the paper reasonably self-contained, we give a brief outline of the basic ideas and refer to [8] for a detailed exposition of the subject. The basic idea is to consider nonlinear submanifolds $\Sigma_N \subset \Sigma_\Lambda$ of the original basis set. These submanifolds

$$\Sigma_N := \left\{ \sum\nolimits_{\lambda \in \Delta} c_\lambda \chi_\lambda : \Delta \subset \Lambda, \#\Delta \leq N \right\}, \tag{2.1}$$

consist of all possible linear combinations of basis functions with at most N terms. For an appropriately chosen separable Hilbert space H, the approximation error with respect to Σ_N is given by

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

For our purposes we consider the Sobolev space H^1 equipped with its natural Hilbert space structure. The approximation spaces of best N-term approximations

$$A_q^\alpha(H) := \{f \in H : |f|_{A_q^\alpha(H)} < \infty\} \qquad \text{with} \quad |f|_{A_q^\alpha(H)} := \sqrt[q]{\sum_{N \in \mathbb{N}} (N^\alpha \sigma_N(f))^q \, N^{-1}}$$

determine the optimal convergence rate that can be achieved with respect to the number of basis functions. Loosely speaking, if $f \in A_q^\alpha(H)$ it is possible to achieve a convergence rate $\sigma_N(f) \sim N^{-\alpha}$. Here and in the following, $a \sim b$ means that the quantities can be uniformly bounded by some constant multiple of each other. Similarly $a \lesssim b$ means that a is uniformly bounded by some constant multiple of b.

A drawback of the initial theory was that it loses some of its efficiency in higher dimensions. Recently, best N-term approximation has been extended by Nitsche [40] to hyperbolic wavelet bases (cf. [10]), which enable an effective treatment of high dimensional problems. This approach is based on hierarchical tensor products that are well known in finite element methods as sparse grids (cf. [1]). In the previous paper [14] we have studied best N-term approximations for one-electron reduced density matrices. For this we considered the fully anisotropic wavelet tensor product basis

$$\tilde{\chi}_{\mathbf{j},\mathbf{a}}(\mathbf{x}) = \psi_{j_1,a_1}(x_1) \, \psi_{j_2,a_2}(x_2) \ldots \psi_{j_6,a_6}(x_6) \qquad \text{with} \quad \begin{cases} \mathbf{j} := (j_1, j_2, \ldots, j_6), \\ \mathbf{a} := (a_1, a_2, \ldots, a_6), \end{cases} \tag{2.2}$$

constructed from orthogonal univariate wavelets $\psi_{j,a}(x) := 2^{j/2}\psi(2^j x - a)$. It has been argued in [12, 13] on the basis of physical considerations that the following hybrid approach might be more appropriate for the description of electron correlations. First of all, multivariate isotropic 3d-wavelets

$$\gamma_{j,\mathbf{a}}^{(\mathbf{s})}(\mathbf{x}) = \psi_{j,a_1}^{(s_1)}(x_1) \, \psi_{j,a_2}^{(s_2)}(x_2) \psi_{j,a_3}^{(s_3)}(x_3) \qquad \text{with} \quad \mathbf{s} := (s_1, s_2, s_3), \ \mathbf{a} := (a_1, a_2, a_3), \tag{2.3}$$

are obtained by taking mixed tensor products of orthogonal univariate wavelets $\psi_{j,a}^{(1)}(x) := 2^{j/2}\psi(2^j x - a)$ and scaling functions $\psi_{j,a}^{(0)}(x) := 2^{j/2}\varphi(2^j x - a)$ on the same level of refinement j. Pure scaling function tensor products $\gamma_{j_0,\mathbf{a}}^{(\mathbf{0})}$ are included on the coarsest level j_0 only. From these wavelets we construct a hybrid anisotropic 6d-wavelet basis

$$\chi_{j_1,j_2,\mathbf{a}_1,\mathbf{a}_2}^{(\mathbf{s}_1,\mathbf{s}_2)}(\mathbf{x},\mathbf{y}) = \gamma_{j_1,\mathbf{a}_1}^{(\mathbf{s}_1)}(\mathbf{x}) \, \gamma_{j_2,\mathbf{a}_2}^{(\mathbf{s}_2)}(\mathbf{y}). \tag{2.4}$$

In Section 2.3 below, we present some rigorous mathematical arguments in favour of this construction. Following Nitsche [40], we consider tensor product Besov spaces

$$\tilde{\mathcal{B}}_q^\alpha(\Omega \times \Omega) = \left\{ B_q^{\alpha+1}(L_q(\Omega)) \bigotimes B_q^\alpha(L_q(\Omega)) \right\} \bigcap \left\{ B_q^\alpha(L_q(\Omega)) \bigotimes B_q^{\alpha+1}(L_q(\Omega)) \right\}$$

for bounded domains $\Omega \subset \mathbb{R}^3$. These spaces are norm equivalent to weighted ℓ_q norms for anisotropic wavelet coefficients

$$\|f\|_{\tilde{\mathcal{B}}_q^\alpha}^q = \sum_{j_1,j_2 \geq j_0} 2^{\max\{j_1,j_2\}q} \left(\sum_{\mathbf{s}_1,\mathbf{s}_2} \sum_{\mathbf{a}_1,\mathbf{a}_2} \left| \langle \chi_{j_1,j_2,\mathbf{a}_1,\mathbf{a}_2}^{(\mathbf{s}_1,\mathbf{s}_2)}, f \rangle \right|^q \right), \qquad \text{if} \quad \alpha = \frac{3}{q} - \frac{3}{2}, \tag{2.5}$$

$$\text{with} \qquad \langle \chi_{j_1,j_2,\mathbf{a}_1,\mathbf{a}_2}^{(\mathbf{s}_1,\mathbf{s}_2)}, f \rangle := \int_{\mathbb{R}^3 \times \mathbb{R}^3} \gamma_{j_1,\mathbf{a}_1}^{(\mathbf{s}_1)}(\mathbf{x}) \, f(\mathbf{x},\mathbf{y}) \, \gamma_{j_2,\mathbf{a}_2}^{(\mathbf{s}_2)}(\mathbf{y}) \, d\mathbf{x} d\mathbf{y}.$$

The norm equivalence requires a univariate wavelet ψ with $p > \alpha + 1$ vanishing moments and $\psi \in B_q^\beta(L_q)$ for some $\beta > \alpha + 1$. The corresponding relation between best N-term approximation spaces and Besov spaces is given by

$$A_q^{\alpha/3}(H^1(\Omega \times \Omega)) = \tilde{\mathcal{B}}_q^\alpha(\Omega \times \Omega), \qquad \text{if} \quad \alpha = \frac{3}{q} - \frac{3}{2}. \tag{2.6}$$

2.1 Besov regularity of Jastrow factors at the $e - e$ cusp

We assume that Jastrow factors possess the asymptotic smoothness property (1.4) near the $e - e$ cusp. If this is the case we can prove the following lemma.

Lemma 2.1 *Suppose $\mathcal{F}^{(2)}$ satisfies Assumption 1.1. Then $\mathcal{F}^{(2)} \in \tilde{\mathcal{B}}_q^\alpha(\Omega \times \Omega)$ for $q > 1$ and $\alpha = \frac{3}{q} - \frac{3}{2}$.*

Proof. Suppose that the isotropic 3d-wavelets $\gamma_{j,\mathbf{a}}^{(\mathbf{s})}$ are constructed from an univariate wavelet ψ with $p > \alpha + 1$ vanishing moments. For each $\gamma_{j,\mathbf{a}}^{(\mathbf{s})}$, we define a cube $\square_{j,\mathbf{a}}$ centred at $2^{-j}\mathbf{a}$ with edge length $2^{-j}L$, such that $\operatorname{supp} \gamma_{j,\mathbf{a}}^{(\mathbf{s})} \subset \square_{j,\mathbf{a}}$. In order to estimate the norm (2.5) for two-particle correlation functions, we restrict ourselves to wavelet coefficients with $j_1 \geq j_2$ and $|\mathbf{s}_1| = |\mathbf{s}_2| = 1$. This combination of 3d-wavelet types corresponds to the worst case where vanishing moments can act in one direction only.

We first consider the case $\operatorname{dist}(\square_{j_1,\mathbf{a}_1}, \square_{j_2,\mathbf{a}_2}) \leq 2^{-j_2}L$. In order to apply the asymptotic smoothness property (1.4), we decompose the cube $\square_{j_2,\mathbf{a}_2}$ into non-overlapping subcubes $\square_i$ ($i \in \Delta$) with edge length $2^{-j_1}L$. The subcubes $\square_i$ with $i \in \Delta_0 := \{i \in \Delta : \operatorname{dist}(\square_{j_1,\mathbf{a}_1}, \square_i) \leq 2^{-j_1}L\}$ are considered separately. Their number is $\#\Delta_0 = O(1)$ independent of the wavelet levels j_1, j_2. For the remaining subcubes $\square_i$ ($i \in \Delta \setminus \Delta_0$) it becomes necessary to control their contributions with respect to $\operatorname{dist}(\square_{j_1,\mathbf{a}_1}, \square_i)$ because $\#(\Delta \setminus \Delta_0) = O(2^{3(j_1-j_2)})$ depends on the wavelet levels. The wavelet coefficients can be estimated by the separate sums

$$\left| \langle \chi_{j_1,j_2,\mathbf{a}_1,\mathbf{a}_2}^{(\mathbf{s}_1,\mathbf{s}_2)}, \mathcal{F}^{(2)} \rangle \right| \leq \sum_{i \in \Delta_0} \left| \int_{\square_{j_1,\mathbf{a}_1} \times \square_i} \gamma_{j_1,\mathbf{a}_1}^{(\mathbf{s}_1)}(\mathbf{x}) \, \mathcal{F}^{(2)}(\mathbf{x},\mathbf{y}) \, \gamma_{j_2,\mathbf{a}_2}^{(\mathbf{s}_2)}(\mathbf{y}) d\mathbf{x}d\mathbf{y} \right| \tag{2.7}$$

$$+ \sum_{i \in \Delta \setminus \Delta_0} \left| \int_{\square_{j_1,\mathbf{a}_1} \times \square_i} \gamma_{j_1,\mathbf{a}_1}^{(\mathbf{s}_1)}(\mathbf{x}) \, \mathcal{F}^{(2)}(\mathbf{x},\mathbf{y}) \, \gamma_{j_2,\mathbf{a}_2}^{(\mathbf{s}_2)}(\mathbf{y}) d\mathbf{x}d\mathbf{y} \right|.$$

For the first sum we can use the following proposition (see, e.g., [14] for details):

Proposition 2.2 *If $f \in Lip(\mathbb{R})$, then*

$$\left| \int_{\mathbb{R}} f(x) \psi_{j,k}^{(1)}(x) dx \right| \lesssim 2^{-3j/2} \|f'\|_{L_\infty(\operatorname{supp} \psi_{j,k}^{(1)})}.$$

Suppose $s_{1,1} = 1$ for the wavelet $\gamma_{j_1,\mathbf{a}_1}^{(\mathbf{s}_1)}$ (see definition (2.3)), we obtain from Proposition 2.2 that

$$\sum_{i \in \Delta_0} \left| \int_{\square_{j_1,\mathbf{a}_1} \times \square_i} \gamma_{j_1,\mathbf{a}_1}^{(\mathbf{s}_1)}(\mathbf{x}) \, \mathcal{F}^{(2)}(\mathbf{x},\mathbf{y}) \, \gamma_{j_2,\mathbf{a}_2}^{(\mathbf{s}_2)}(\mathbf{y}) d\mathbf{x}d\mathbf{y} \right| \tag{2.8}$$

$$\lesssim \underbrace{2^{-\frac{3}{2}j_1}}_{(a)} \underbrace{2^{-j_1}}_{(b)} \underbrace{2^{-3j_1} 2^{\frac{3}{2}j_2}}_{(c)} \|\partial_{x_1} \mathcal{F}^{(2)}\|_{L_\infty(\square_{j_1,\mathbf{a}_1} \times \square_i)} \lesssim 2^{-\frac{11}{2}j_1} 2^{\frac{3}{2}j_2},$$

where the prefactors constitute in the following way: (a) Proposition 2.2 for $\psi_{j_1,a_1}^{(1)}(x_1)$, (b) normalization constant of $\psi_{j_1,a_2}^{(s_{1,2})}(x_2)\psi_{j_1,a_3}^{(s_{1,3})}(x_3)$ and sectional area of $\square_{j_1,\mathbf{a}_1}$ in the x_2, x_3 plane, (c) volume of the cube $\square_i$ and normalization constant of $\gamma_{j_2,\mathbf{a}_2}^{(\mathbf{s}_2)}$.

The second sum can be estimated using the next proposition (see, e.g., [14] for details):

Proposition 2.3 *Suppose the function $f(\mathbf{x})$ with $\mathbf{x} \in \mathbb{R}^3$ is smooth on the support of an isotropic 3d-wavelet $\gamma_{j,\mathbf{a}}^{(\mathbf{s})}$. Then the following estimate holds,*

$$\left| \int_{\mathbb{R}^3} f(\mathbf{x}) \, \gamma_{j,\mathbf{a}}^{(\mathbf{s})}(\mathbf{x}) d\mathbf{x} \right| \lesssim 2^{-(p|\mathbf{s}|+3/2)j} \|\partial_{x_1}^{s_1 p} \partial_{x_2}^{s_2 p} \partial_{x_3}^{s_3 p} f\|_{L_\infty(\operatorname{supp} \gamma_{j,\mathbf{a}}^{(\mathbf{s})})} \quad with \; |\mathbf{s}| := s_1 + s_2 + s_3.$$

With this and the estimate (1.4) for wavelets $\gamma_{j_1,\mathbf{a}_1}^{(\mathbf{s}_1)}$ with p vanishing moments (i.e., $|\mathbf{s}_1| = 1$), we obtain

$$\sum_{i \in \Delta \backslash \Delta_0} \left| \int_{\square_{j_1,\mathbf{a}_1} \times \square_i} \gamma_{j_1,\mathbf{a}_1}^{(\mathbf{s}_1)}(\mathbf{x})\, \mathcal{F}^{(2)}(\mathbf{x},\mathbf{y})\, \gamma_{j_2,\mathbf{a}_2}^{(\mathbf{s}_2)}(\mathbf{y}) d\mathbf{x} d\mathbf{y} \right| \lesssim \sum_{i \in \Delta \backslash \Delta_0} \underbrace{2^{-(p+\frac{3}{2})j_1}}_{(a)} \underbrace{2^{-3j_1} 2^{\frac{3}{2}j_2}}_{(b)} \|\partial_{x_1}^p \mathcal{F}^{(2)}\|_{L_\infty(\square_{j_1,\mathbf{a}_1} \times \square_i)}$$

$$\lesssim 2^{-(p+\frac{3}{2})j_1} 2^{\frac{3}{2}j_2} \int_{2^{-j_1}L}^{2^{-j_2+2}L} r^{3-p}\, dr \lesssim \begin{cases} 2^{-(p+\frac{3}{2})j_1} 2^{-(\frac{5}{2}-p)j_2} & \text{if } p \leq 3 \\ (j_1 - j_2 + 1) 2^{-\frac{11}{2}j_1} 2^{\frac{3}{2}j_2} & \text{if } p = 4 \\ 2^{-\frac{11}{2}j_1} 2^{\frac{3}{2}j_2} & \text{if } p > 4 \end{cases}, \tag{2.9}$$

where the prefactor in the second line constitutes in the following way: (a) Proposition 2.3 for $\gamma_{j_1,\mathbf{a}_1}^{(\mathbf{s}_1)}$, (b) volume of the cube $\square_i$ and normalization constant of $\gamma_{j_2,\mathbf{a}_2}^{(\mathbf{s}_2)}$.

Once we have obtained the estimates (2.8) and (2.9), it is straightforward to get an upper bound for the contribution of anisotropic tensor products with translation parameters $(\mathbf{a}_1,\mathbf{a}_2) \in A_{j_1,j_2} := \{(\mathbf{a}_1,\mathbf{a}_2) : \text{dist}(\square_{j_1,\mathbf{a}_1}, \square_{j_2,\mathbf{a}_2}) \leq 2^{-j_2}L\}$ to the norm (2.5),

$$\sum_{j_1 \geq j_2 \geq j_0} 2^{qj_1} \sum_{(\mathbf{a}_1,\mathbf{a}_2) \in A_{j_1,j_2}} \left| \langle \chi_{j_1,j_2,\mathbf{a}_1,\mathbf{a}_2}^{(\mathbf{s}_1,\mathbf{s}_2)}, \mathcal{F}^{(2)} \rangle \right|^q \tag{2.10}$$

$$\lesssim \sum_{j_1 \geq j_2 \geq j_0} \begin{cases} 2^{-(pq+\frac{9}{2}-3)j_1} 2^{-(\frac{5}{2}-p)qj_2} & \text{if } p \leq 3 \\ (j_1 - j_2 + 1) 2^{-(\frac{9}{2}q-3)j_1} 2^{\frac{3}{2}qj_2} & \text{if } p = 4 \\ 2^{-(\frac{9}{2}q-3)j_1} 2^{\frac{3}{2}qj_2} & \text{if } p > 4 \end{cases} < \infty, \quad \text{if } \begin{cases} q > \frac{6}{5} & \text{for } p = 2, \\ q > 1 & \text{for } p \geq 3, \end{cases}$$

where we have used $\# A_{j_1,j_2} = O(2^{3j_1})$.

In order to get an upper bound for the norm (2.5) it remains to estimate the contributions of anisotropic wavelet coefficients where the supports of the corresponding 3d-wavelets are well separated. For this we have to consider the parameter set $B_{j_1,j_2} := \{(\mathbf{a}_1,\mathbf{a}_2) : 2^{-j_2}L < \text{dist}(\square_{j_1,\mathbf{a}_1}, \square_{j_2,\mathbf{a}_2})\}$. Using estimate (1.4) and Proposition 2.3, the contributions can be estimated

$$\sum_{(\mathbf{a}_1,\mathbf{a}_2) \in B_{j_1,j_2}} \left| \langle \chi_{j_1,j_2,\mathbf{a}_1,\mathbf{a}_2}^{(\mathbf{s}_1,\mathbf{s}_2)}, \mathcal{F}^{(2)} \rangle \right|^q \lesssim \sum_{(\mathbf{a}_1,\mathbf{a}_2) \in B_{j_1,j_2}} 2^{-(p+\frac{3}{2})q(j_1+j_2)} \|\partial_{x_1}^p \partial_{y_1}^p \mathcal{F}^{(2)}\|_{L_\infty(\square_{j_1,\mathbf{a}_1} \times \square_{j_2,\mathbf{a}_2})}^q$$

$$\lesssim 2^{-(pq+\frac{3}{2}q-3)j_1} 2^{-(pq+\frac{3}{2}q-3)j_2} \int_{2^{-j_2}L}^{\text{diam}\,\Omega} r^{2+q(1-2p)}\, dr \lesssim 2^{-(qp+\frac{3}{2}q-3)j_1} 2^{(p-\frac{5}{2})qj_2},$$

where we have used $p > \alpha + 1$ and $\alpha = \frac{3}{q} - \frac{3}{2}$, hence

$$2 - 2pq + q < -1 - q\left(p - \frac{3}{2}\right) < -1, \qquad \text{since } p \geq 2,$$

follows for the exponent of the integrand. The remaining sum with respect to the wavelet levels yields

$$\sum_{j_1 \geq j_2 \geq j_0} 2^{qj_1} \sum_{(\mathbf{a}_1,\mathbf{a}_2) \in B_{j_1,j_2}} \left| \langle \chi_{j_1,j_2,\mathbf{a}_1,\mathbf{a}_2}^{(\mathbf{s}_1,\mathbf{s}_2)}, \mathcal{F}^{(2)} \rangle \right|^q \lesssim \sum_{j_1 \geq j_2 \geq j_0} 2^{-(qp+\frac{1}{2}q-3)j_1} 2^{(p-\frac{5}{2})qj_2}$$

$$\lesssim \sum_{j_1 \geq j_0} \begin{cases} 2^{-(\frac{5}{2}q-3)j_1} & \text{if } p = 2 \\ 2^{-3(q-1)j_1} & \text{if } p \geq 3 \end{cases} < \infty, \quad \text{if } \begin{cases} q > \frac{6}{5} & \text{for } p = 2, \\ q > 1 & \text{for } p \geq 3, \end{cases} \tag{2.11}$$

from which we obtain, together with our previous estimate (2.10), the lower bound on the Besov space parameter q. ∎

Lemma 2.1 demonstrates that there exists an upper bound on the regularity of two-particle correlation functions $\mathcal{F}^{(2)}$ in tensor product Besov spaces $\tilde{\mathcal{B}}_q^\alpha$ and therefore, via relation (2.6), restrictions on the best N-term approximation spaces $A_q^{\alpha/3}(H^1(\Omega \times \Omega))$ to which they belong. Only in the special case of wavelets with $p = 2$ vanishing moments, the function $\mathcal{F}^{(2)}$ belongs to all best N-term approximation spaces that are accessible for such kind of wavelet bases. The following corollary shows that the bound of Lemma 2.1 has to be considered as sharp. For this we consider an explicit example which represents, up to a constant prefactor, the simplest possible two-particle correlation function consistent with Kato's [31] cusp condition.

Corollary 2.4 *The two-particle correlation function* $\mathcal{F}_c^{(2)}(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$ *satisfies Assumption 1.1 and does not belong to* $\tilde{\mathcal{B}}_q^\alpha(\Omega \times \Omega)$ *for* $q \leq 1$ *and* $\alpha = \frac{3}{q} - \frac{3}{2}$.

Proof. We consider the norm equivalence (2.5) and restrict ourselves to the contributions of diagonal tensor products

$$\|\mathcal{F}_c^{(2)}\|_{\tilde{\mathcal{B}}_q^\alpha}^q \geq \sum_{j \geq j_0} 2^{jq} \sum_{\mathbf{s}} \sum_{\mathbf{a}} \left| \langle \chi_{j,j,\mathbf{a},\mathbf{a}}^{(\mathbf{s},\mathbf{s})}, \mathcal{F}_c^{(2)} \rangle \right|^q \tag{2.12}$$

with

$$\langle \chi_{j,j,\mathbf{a},\mathbf{a}}^{(\mathbf{s},\mathbf{s})}, \mathcal{F}_c^{(2)} \rangle := \int_{\mathbb{R}^3 \times \mathbb{R}^3} \gamma_{j,\mathbf{a}}^{(\mathbf{s})}(\mathbf{x}) \, |\mathbf{x} - \mathbf{y}| \, \gamma_{j,\mathbf{a}}^{(\mathbf{s})}(\mathbf{y}) \, d\mathbf{x} d\mathbf{y}.$$

Using the scaling relation

$$\langle \chi_{j,j,\mathbf{a},\mathbf{a}}^{(\mathbf{s},\mathbf{s})}, \mathcal{F}_c^{(2)} \rangle = 2^{-4j} \langle \chi_{0,0,\mathbf{0},\mathbf{0}}^{(\mathbf{s},\mathbf{s})}, \mathcal{F}_c^{(2)} \rangle$$

for the wavelet coefficients, we obtain for the sum

$$\sum_{\mathbf{a}} \left| \langle \chi_{j,j,\mathbf{a},\mathbf{a}}^{(\mathbf{s},\mathbf{s})}, \mathcal{F}_c^{(2)} \rangle \right|^q = C \, 2^{3j} 2^{-4jq} \left| \langle \chi_{0,0,\mathbf{0},\mathbf{0}}^{(\mathbf{s},\mathbf{s})}, \mathcal{F}_c^{(2)} \rangle \right|^q. \tag{2.13}$$

Inserting Eq. (2.13) into the estimate (2.12), we get $\|\mathcal{F}_c^{(2)}\|_{\tilde{\mathcal{B}}_q^\alpha}^q = \infty$ for $q \leq 1$. ∎

2.2 Besov regularity of Jastrow factors with off-diagonal electron-nuclear cusp

Lemma 2.5 *Suppose* $\mathcal{F}^{(2)}$ *satisfies Assumption 1.2. Then* $\mathcal{F}^{(2)} \in \tilde{\mathcal{B}}_q^\alpha(\Omega \times \Omega)$ *for* $q > 1$ *and* $\alpha = \frac{3}{q} - \frac{3}{2}$.

Proof. For the sake of notational simplicity, we assume without loss of generality that $\mathbf{R} = \mathbf{0}$. Furthermore it is sufficient to consider the case of $p > 4$ vanishing moments. Like for the proof of Lemma 2.1 it is possible to perform all estimates with $p > \alpha + 1$ vanishing moments. This would not, however, provide any new insight and instead makes the whole proof unnecessarily long-winded. At first, we provide modified estimates for the partial derivatives of $\mathcal{F}^{(2)}$ which are more convenient from an analytical point of view. The following estimates are an immediate consequence of the basic Assumption 1.2,

$$|\partial_\mathbf{x}^\alpha \partial_\mathbf{y}^\beta \mathcal{F}^{(2)}(\mathbf{x}, \mathbf{y})| \lesssim \begin{cases} \left(\delta_{|\alpha|} + |\mathbf{x}|^{1-|\alpha|} \right) \left(\delta_{|\beta|} + |\mathbf{y}|^{1-|\beta|} \right) & \text{if } \min\{|\mathbf{x}|, |\mathbf{y}|\} < |\mathbf{x} - \mathbf{y}| \\ |\mathbf{x} - \mathbf{y}|^{1-|\alpha|-|\beta|} & \text{if } |\mathbf{x} - \mathbf{y}| \leq |\mathbf{x}|, |\mathbf{y}| \end{cases} \tag{2.14}$$

$$\lesssim \left(\delta_{|\alpha|} + |\mathbf{x}|^{1-|\alpha|} \right) \left(\delta_{|\beta|} + |\mathbf{y}|^{1-|\beta|} \right) + |\mathbf{x} - \mathbf{y}|^{1-|\alpha|-|\beta|}$$

for $(\mathbf{x}, \mathbf{y}) \notin D \cup N$ and $|\alpha| + |\beta| > 0$.

In the lemma, we consider four different cases with respect to the spatial configurations of the hypercubes $\square_{j_1, \mathbf{a}_1} \times \square_{j_2, \mathbf{a}_2}$ which contain the supports of tensor product wavelets $\chi_{j_1, j_2, \mathbf{a}_1, \mathbf{a}_2}^{(\mathbf{s}_1, \mathbf{s}_2)}$.

Case a:

First we consider the set of parameters

$$A_{j_1, j_2} := \{(\mathbf{a}_1, \mathbf{a}_2) : \text{dist}(\square_{j_1, \mathbf{a}_1}, \mathbf{0}) \leq 2^{-j_1+\epsilon} L, \, \text{dist}(\square_{j_2, \mathbf{a}_2}, \mathbf{0}) \leq 2^{-j_2+\epsilon} L\}$$

with cardinality $\#A_{j_1, j_2} = O(1)$. The parameter $\epsilon > 0$ will be fixed in the sequel. Let us assume $j_1 \geq j_2$. Using Lipschitz continuity of $\mathcal{F}^{(2)}$, we derive from Proposition 2.2 the following estimate for the wavelet coefficients:

$$\left| \langle \chi_{j_1, j_2, \mathbf{a}_1, \mathbf{a}_2}^{(\mathbf{s}_1, \mathbf{s}_2)}, \mathcal{F}^{(2)} \rangle \right| \lesssim 2^{-\frac{5}{2} j_1} 2^{-\frac{3}{2} j_2} \|\partial_{x_1} \mathcal{F}^{(2)}\|_{L_\infty(\square_{j_1, \mathbf{a}_1} \times \square_{j_2, \mathbf{a}_2})} \lesssim 2^{-\frac{5}{2} j_1} 2^{-\frac{3}{2} j_2}. \tag{2.15}$$

This immediately shows the boundedness of the contributions from A_{j_1, j_2} to the Besov norm (2.5):

$$\sum_{j_1 \geq j_2 \geq j_0} 2^{j_1 q} \sum_{(\mathbf{a}_1, \mathbf{a}_2) \in A_{j_1, j_2}} \left| \langle \chi_{j_1, j_2, \mathbf{a}_1, \mathbf{a}_2}^{(\mathbf{s}_1, \mathbf{s}_2)}, \mathcal{F}^{(2)} \rangle \right|^q \lesssim \sum_{j_1 \geq j_2 \geq j_0} 2^{-\frac{3}{2} q(j_1 + j_2)} < \infty.$$

Case b:

Here we consider the set of parameters $B_{j_1,j_2} \cap C_{j_1,j_2}$,

$$B_{j_1,j_2} := \{(\mathbf{a}_1,\mathbf{a}_2) : \operatorname{dist}(\square_{j_1,\mathbf{a}_1},\mathbf{0}) \le 2^{-j_1+\epsilon}L,\ \operatorname{dist}(\square_{j_2,\mathbf{a}_2},\mathbf{0}) > 2^{-j_2+\epsilon}L\},$$
$$C_{j_1,j_2} := \{(\mathbf{a}_1,\mathbf{a}_2) : \operatorname{dist}(\square_{j_1,\mathbf{a}_1},\square_{j_2,\mathbf{a}_2}) \le 2^{-\min\{j_1,j_2\}}L\}.$$

For $j_1 \ge j_2$ we can use the previous estimate (2.15) because of $\#(B_{j_1,j_2} \cap C_{j_1,j_2}) = O(1)$. The remaining combination $j_2 > j_1$ requires a decomposition of the form (2.7) for the cube $\square_{j_1,\mathbf{a}_1}$. Using the estimates (2.8), (2.9), and (2.14) we obtain for the wavelet coefficients

$$\left| \langle \chi_{j_1,j_2,\mathbf{a}_1,\mathbf{a}_2}^{(\mathbf{s}_1,\mathbf{s}_2)}, \mathcal{F}^{(2)} \rangle \right| \lesssim 2^{\frac{3}{2}j_1} 2^{-\frac{11}{2}j_2} + \sum_{i \in \Delta \setminus \Delta_0} \left| \int_{\square_i \times \square_{j_2,\mathbf{a}_2}} \gamma_{j_1,\mathbf{a}_1}^{(\mathbf{s}_1)}(\mathbf{x})\, \mathcal{F}^{(2)}(\mathbf{x},\mathbf{y})\, \gamma_{j_2,\mathbf{a}_2}^{(\mathbf{s}_2)}(\mathbf{y}) d\mathbf{x} d\mathbf{y} \right|$$

$$\lesssim 2^{\frac{3}{2}j_1} 2^{-\frac{11}{2}j_2} + \sum_{i \in \Delta \setminus \Delta_0} 2^{\frac{3}{2}j_1} 2^{-(p+\frac{9}{2})j_2} \|\partial_{y_1}^p \mathcal{F}^{(2)}\|_{L_\infty(\square_i \times \square_{j_2,\mathbf{a}_2})}$$

$$\lesssim 2^{\frac{3}{2}j_1} 2^{-\frac{11}{2}j_2} + 2^{-\frac{3}{2}j_1} 2^{-(p+\frac{3}{2})j_2} |2^{-j_2}\mathbf{a}_2|^{1-p},$$

where $\#\{\square_i\} = O(2^{3(j_2-j_1)})$ has been used in the last line. With this and $\#(B_{j_1,j_2} \cap C_{j_1,j_2}) = O(2^{3(j_2-j_1)})$, we can estimate the contributions from $B_{j_1,j_2} \cap C_{j_1,j_2}$ to the Besov norm (2.5),

$$\sum_{j_2 \ge j_1 \ge j_0} 2^{j_2 q} \sum_{(\mathbf{a}_1,\mathbf{a}_2) \in B_{j_1,j_2} \cap C_{j_1,j_2}} \left| \langle \chi_{j_1,j_2,\mathbf{a}_1,\mathbf{a}_2}^{(\mathbf{s}_1,\mathbf{s}_2)}, \mathcal{F}^{(2)} \rangle \right|^q$$

$$\lesssim \sum_{j_2 \ge j_1 \ge j_0} \left(2^{(\frac{3}{2}q-3)j_1} 2^{-(\frac{9}{2}q-3)j_2} + 2^{-\frac{3}{2}qj_1} 2^{-q(p+\frac{1}{2})j_2} \sum_{\mathbf{a}_2} |2^{-j_2}\mathbf{a}_2|^{q(1-p)} \right)$$

$$\lesssim \sum_{j_2 \ge j_1 \ge j_0} \left(2^{-3j_1} 2^{-3(q-1)j_2} + 2^{-\frac{3}{2}qj_1} 2^{-(qp+\frac{q}{2}-3)j_2} \int_{2^{-j_2+\epsilon}L}^\infty r^{2+q(1-p)} dr \right)$$

$$\lesssim \sum_{j_2 \ge j_1 \ge j_0} \left(2^{-3j_1} 2^{-3(q-1)j_2} + 2^{-\frac{3}{2}q(j_1+j_2)} \right) < \infty \qquad \text{since } q > 1.$$

Case c:

We consider now the parameter set $B_{j_1,j_2} \setminus C_{j_1,j_2}$, where

$$|\mathbf{y}| \lesssim |\mathbf{x}-\mathbf{y}| \qquad \text{for} \quad (\mathbf{x},\mathbf{y}) \in \square_{j_1,\mathbf{a}_1} \times \square_{j_2,\mathbf{a}_2}.$$

The wavelet coefficients can be estimated by

$$\left| \langle \chi_{j_1,j_2,\mathbf{a}_1,\mathbf{a}_2}^{(\mathbf{s}_1,\mathbf{s}_2)}, \mathcal{F}^{(2)} \rangle \right| \lesssim 2^{-\frac{3}{2}j_1} 2^{-(p+\frac{3}{2})j_2} \|\partial_{y_1}^p \mathcal{F}^{(2)}\|_{L_\infty(\square_{j_1,\mathbf{a}_1} \times \square_{j_2,\mathbf{a}_2})} \lesssim 2^{-\frac{3}{2}j_1} 2^{-(p+\frac{3}{2})j_2} |2^{-j_2}\mathbf{a}_2|^{1-p}.$$

From this, the boundedness of the contributions from the parameter set $B_{j_1,j_2} \setminus C_{j_1,j_2}$ to the Besov norm (2.5) follows through the estimate

$$\sum_{j_1,j_2 \ge j_0} 2^{\max\{j_1,j_2\}q} \sum_{(\mathbf{a}_1,\mathbf{a}_2) \in B_{j_1,j_2} \setminus C_{j_1,j_2}} \left| \langle \chi_{j_1,j_2,\mathbf{a}_1,\mathbf{a}_2}^{(\mathbf{s}_1,\mathbf{s}_2)}, \mathcal{F}^{(2)} \rangle \right|^q$$

$$\lesssim \sum_{j_1,j_2 \ge j_0} 2^{\max\{j_1,j_2\}q} 2^{-\frac{3}{2}qj_1} 2^{-(qp+\frac{3}{2}q-3)j_2} \int_{2^{-j_2+\epsilon}L}^\infty r^{2+q(1-p)}\, dr$$

$$\lesssim \sum_{j_1,j_2 \ge j_0} 2^{\max\{j_1,j_2\}q} 2^{-\frac{3}{2}qj_1} 2^{-\frac{5}{2}qj_2} < \infty.$$

Case d:

Let us assume $j_1 \ge j_2$, and consider first the set of parameters $D_{j_1,j_2} \cap C_{j_1,j_2}$ with

$$D_{j_1,j_2} := \{(\mathbf{a}_1,\mathbf{a}_2) : 2^{-j_1+\epsilon}L < \operatorname{dist}(\square_{j_1,\mathbf{a}_1},\mathbf{0}),\ 2^{-j_2+\epsilon}L < \operatorname{dist}(\square_{j_2,\mathbf{a}_2},\mathbf{0})\},$$

9

where we can choose the parameter ϵ in such a way that

$$|\mathbf{x} - \mathbf{y}| \lesssim |\mathbf{x}|, |\mathbf{y}| \qquad \text{for} \quad (\mathbf{x}, \mathbf{y}) \in \square_{j_1, \mathbf{a}_1} \times \square_{j_2, \mathbf{a}_2}.$$

Due to this relation, we can use the estimates (2.8), (2.9) from the first part of the proof of Lemma 2.1 in order to show the boundedness of this contribution to the Besov norm (2.5).

It remains to consider the set of parameters $D_{j_1, j_2} \setminus C_{j_1, j_2}$:

$$\left| \langle \chi^{(\mathbf{s}_1, \mathbf{s}_2)}_{j_1, j_2, \mathbf{a}_1, \mathbf{a}_2}, \mathcal{F}^{(2)} \rangle \right|^q \lesssim 2^{-(p+\frac{3}{2})q(j_1+j_2)} \| \partial_{x_1}^p \partial_{y_1}^p \mathcal{F}^{(2)} \|^q_{L^\infty(\square_{j_1,\mathbf{a}_1} \times \square_{j_2,\mathbf{a}_2})}$$
$$\lesssim 2^{-(p+\frac{3}{2})q(j_1+j_2)} \left(|2^{-j_1}\mathbf{a}_1 - 2^{-j_2}\mathbf{a}_2|^{q(1-2p)} + |2^{-j_1}\mathbf{a}_1|^{q(1-p)} |2^{-j_2}\mathbf{a}_2|^{q(1-p)} \right).$$

For the first term, we can use estimate (2.11) from the second part of the proof of Lemma 2.1. The contributions from the second term are also bounded according to the following estimate

$$\sum_{j_1 \geq j_2 \geq j_0} 2^{j_1 q} 2^{-(p+\frac{3}{2})q(j_1+j_2)} \sum_{(\mathbf{a}_1, \mathbf{a}_2) \in D_{j_1, j_2} \setminus C_{j_1, j_2}} |2^{-j_1}\mathbf{a}_1|^{q(1-p)} |2^{-j_2}\mathbf{a}_2|^{q(1-p)}$$
$$\lesssim \sum_{j_1 \geq j_2 \geq j_0} 2^{-(qp+\frac{q}{2}-3)j_1} 2^{-(qp+\frac{3}{2}q-3)j_2} \int_{2^{-j_1+\epsilon}L}^{\infty} r_1^{2+q(1-p)} \, dr_1 \int_{2^{-j_2+\epsilon}L}^{\infty} r_2^{2+q(1-p)} \, dr_2$$
$$\lesssim \sum_{j_1 \geq j_2 \geq j_0} 2^{-\frac{3}{2}qj_1} 2^{-\frac{5}{2}qj_2} < \infty.$$

Altogether, this demonstrates the finiteness of the Besov space norm (2.5) for $\mathcal{F}^{(2)}$. ∎

2.3 Practical implications for electronic structure calculations

It is an immediate consequence of Lemmata 2.1 and 2.5 that there exists a sharp upper bound on the approximation error of best N-term approximations for two-particle correlation functions

$$\sigma_N(\mathcal{F}^{(2)}) := \inf_{\mathcal{F}^{(2)}_N \in \Sigma_N} \| \mathcal{F}^{(2)} - \mathcal{F}^{(2)}_N \|_{H^1} \lesssim N^{-\frac{1}{2}} \tag{2.16}$$

where $\mathcal{F}^{(2)}_N$ belongs to the nonlinear submanifold (2.1), irrespectively of the number of vanishing moments $p \geq 3$ of the wavelet basis. It is not surprising with respect to our previous result for the one-electron reduced density matrix (cf. [14]) that $e-n$ cusps have no effect on the approximation properties of anisotropic tensor product bases. This has already been observed by Nitsche [39] for boundary value problems with point-like singularities. Concerning the choice of the multivariate wavelet basis it has been argued in [14] that a fully anisotropic basis (2.2) is considerably more favourable for the density matrix compared to the hybrid approach (2.4). For Jastrow factors this is not the case any more. It is easy to see that the upper bound for a fully anisotropic wavelet basis cannot be better than (2.16). However in order to reach this bound it requires only $p \geq 2$ vanishing moments for the underlying univariate wavelet basis compared to $p \geq 3$ vanishing moments for the hybrid approach (2.4). Concerning the computational complexity this is only a minor advantage which has to be compared with the considerably more complicated matrix structure for fully anisotropic wavelets. Furthermore it has been argued in [13] that the vanishing moments property enables the compression of integrals which are required for the construction of Galerkin matrices. Another argument in favour of fairly regular wavelets is the efficient expansion of pointwise wavelet products [12] appearing in these integrals. For such reasons it is desirable to use wavelet bases with a moderate number of vanishing moments (e.g., $p \approx 5$) which meet the regularity requirements of best N-term approximations. According to these arguments one might wonder whether fully isotropic wavelets

$$\breve{\chi}^{(\mathbf{s})}_{j, \mathbf{a}}(\mathbf{x}) = \psi^{(s_1)}_{j, a_1}(x_1) \, \psi^{(s_2)}_{j, a_2}(x_2) \ldots \psi^{(s_6)}_{j, a_6}(x_6) \qquad \text{with} \quad \mathbf{s} := (s_1, s_2, \ldots, s_6), \ \mathbf{a} := (a_1, a_2, \ldots, a_6),$$

where the bound (2.16) requires $p \geq 4$ vanishing moments, are an even better choice. Numerical studies, however, revealed that two-particle correlation functions have a pronounced hyberbolic-cross structure on coarse scales (cf. [13]). This sparsity, which is of considerable practical importance, would be lost for fully isotropic wavelet bases.

3 Comparison with other approximation schemes

3.1 Sparse grid spaces of mixed partial derivatives

Recently it has been proven by Yserentant [50] that exact solutions of the many-electron Schrödinger equation belong to Sobolev spaces of mixed partial derivatives. This result enables the construction of sparse grid approximations for the entire wavefunction [21, 51]. It is instructive to apply this method to Jastrow factors, because Nitsche's [40] approach to best N-term approximation actually corresponds to an adaptive variant of sparse grids. For this we have to consider two-particle correlation functions in mixed Sobolev norms

$$\|\mathcal{F}^{(2)}\|_{H_{\mathrm{mix}}^{s,1}} := \int_{\mathbb{R}^3 \times \mathbb{R}^3} \left(1 + |\mathbf{k}_1|^2 + |\mathbf{k}_2|^2\right) \left(1 + |\mathbf{k}_1|^2\right)^s \left(1 + |\mathbf{k}_2|^2\right)^s |\hat{\mathcal{F}}^{(2)}(\mathbf{k}_1, \mathbf{k}_2)|^2 d\mathbf{k}_1 d\mathbf{k}_2, \ s > 0, \qquad (3.1)$$

$$\text{with} \quad \hat{\mathcal{F}}^{(2)}(\mathbf{k}_1, \mathbf{k}_2) = \int_{\mathbb{R}^3 \times \mathbb{R}^3} \mathcal{F}^{(2)}(\mathbf{x}, \mathbf{y}) e^{-i\mathbf{k}_1 \mathbf{x}} e^{-i\mathbf{k}_2 \mathbf{y}} d\mathbf{x} d\mathbf{y}. \qquad (3.2)$$

Instead of presenting a general treatment of this subject, we restrict ourselves to the illustrative example from Corollary 2.4. In order to allow for mixed Sobolev norms in $\mathbb{R}^3 \times \mathbb{R}^3$, convenient cut-off functions have to be added.

Observation 3.1 *The two-particle correlation function*

$$\mathcal{F}_c^{(2)}(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}| e^{-(|\mathbf{x}|^2 + |\mathbf{y}|^2)}, \qquad (3.3)$$

which satisfies Assumption 1.1, belongs to the Sobolev spaces of mixed partial derivatives $H_{mix}^{s,1}(\mathbb{R}^3 \times \mathbb{R}^3)$, i.e., $\|\mathcal{F}_c^{(2)}\|_{H_{mix}^{s,1}} < \infty$ for all $0 < s < \frac{3}{4}$.

Proof. A straightforward calculation shows that the Fourier transform (3.2) of the two-particle correlation function (3.3) can be represented in terms of an oscillatory integral

$$\hat{\mathcal{F}}_c^{(2)}(\mathbf{k}_1, \mathbf{k}_2) = 2^{\frac{3}{2}} \pi^{\frac{5}{2}} e^{-\frac{1}{8}|\mathbf{k}_1 + \mathbf{k}_2|^2} \frac{1}{|\mathbf{k}_1 - \mathbf{k}_2|} \int_0^\infty s^2 e^{-\frac{1}{2}s^2} \sin\left(\frac{1}{2}|\mathbf{k}_1 - \mathbf{k}_2|s\right) ds. \qquad (3.4)$$

The asymptotic behaviour of oscillatory integrals of this kind is given by the general formula

$$\int_0^\infty e^{isx} f(s) s^n \, ds \sim \sum_{j=0}^\infty a_j x^{-j-n-1} \quad \text{for } n > -1 \text{ with } a_j = i^{j+n+1} \frac{\Gamma(j+n+1)}{j!} f^{(j)}(0), \qquad (3.5)$$

which can be applied to any function $f \in C^\infty(\mathbb{R})$ that vanishes together with all of its derivatives at infinity. For a proof of this formula we refer, e.g., to Copson's monograph [5]. For our present purpose it is sufficient to determine the leading order term of Eq. (3.4),

$$\hat{\mathcal{F}}_c^{(2)}(\mathbf{k}_1, \mathbf{k}_2) \sim 2^{\frac{3}{2}} \pi^{\frac{5}{2}} e^{-\frac{1}{8}|\mathbf{k}_1 + \mathbf{k}_2|^2} \left[-\frac{2^4}{|\mathbf{k}_1 - \mathbf{k}_2|^4} + O\left(\frac{1}{|\mathbf{k}_1 - \mathbf{k}_2|^6}\right) \right] \quad \text{for} \quad |\mathbf{k}_1 - \mathbf{k}_2| \to \infty. \qquad (3.6)$$

Introducing new variables $\mathbf{k}_1 = \mathbf{p} + \mathbf{t}$, $\mathbf{k}_2 = \mathbf{p} - \mathbf{t}$, we define $\tilde{\hat{\mathcal{F}}}_c^{(2)}(\mathbf{p}, \mathbf{t}) := \hat{\mathcal{F}}_c^{(2)}(\mathbf{p} + \mathbf{t}, \mathbf{p} - \mathbf{t})$. From the estimates

$$|\mathbf{k}_i|^{2s} \leq (|\mathbf{p}| + |\mathbf{t}|)^{2s} \leq 2^{2s} \left(|\mathbf{p}|^{2s} + |\mathbf{t}|^{2s}\right) \qquad \text{with } i = 1, 2,$$

we obtain an estimate for the mixed Sobolev norm (3.1)

$$\|\mathcal{F}_c^{(2)}\|_{H_{\mathrm{mix}}^{s,1}} \lesssim \int_{\mathbb{R}^3 \times \mathbb{R}^3} \left(1 + |\mathbf{p}|^2 + |\mathbf{t}|^2\right) \left(1 + |\mathbf{p}|^{2s} + |\mathbf{t}|^{2s}\right)^2 |\tilde{\hat{\mathcal{F}}}_c^{(2)}(\mathbf{p}, \mathbf{t})|^2 d\mathbf{p} d\mathbf{t}$$

in terms of the new variables. Together with Eq. (3.4) and the asymptotic expansion (3.6), the observation is now an obvious consequence. ∎

The sparse-grids approximation scheme for wavefunctions recently suggested by Yserentant [51] can be also applied to the two-particle correlation function. It provides an error estimate

$$\|\mathcal{F}^{(2)} - \mathcal{F}_N^{(2)}\|_{H^1} \lesssim 2^{-Ls} \|\mathcal{F}_c^{(2)}\|_{H_{\mathrm{mix}}^{s,1}}$$

11

with respect to the sparse-grids parameter L. Similarly to the wavelet tensor products (2.4), Yserentant constructed an anisotropic tensor product basis from hierarchical 3d-basis functions. According to the standard approach to sparse grids, this basis contains all tensor products with $\ell_1 + \ell_2 \leq L$, where ℓ_i corresponds to the levels of the underlying 3d-basis functions. The cardinality of the 3d-basis on level ℓ is of $O(2^{3\ell})$. Therefore the cardinality of the tensor product basis increases with $O(L2^{3L})$. This yields a convergence of almost $O(N^{-1/4})$ in the Sobolev space H^1, which is half of the convergence rate (2.16) that can be achieved for best N-term approximation.

3.2 Gaussian-type orbital bases

Approximation of correlated wavefunctions is a central topic in quantum chemistry (cf. [25]) with long-standing history. Despite the tremendous amount of work which has been devoted to the development of many-particle models, like *coupled cluster* (CC) or many-body perturbation theory, only little is known about approximation properties of the underlying discretization schemes. Some rigorous results concerning partial wave expansions of atomic wavefunctions have been derived by Hill-Kutzelnigg-Morgan [27, 35, 36]. The partial wave expansions and atomic centred *Gaussian-type orbital* (GTO) basis sets [25], commonly used in quantum chemistry, share the angular momentum quantum number ℓ as a common ordering parameter. However GTO basis sets have as an additional parameter, namely the exponent of the Gaussians which determines the radial approximation of the basis set. In correlation-consistent GTO basis sets (cf. [25]), usually denoted as cc-pVXZ bases, angular and radial parameters are combined in a balanced manner into a single parameter $X = 2, 3, 4, \ldots$. For these basis sets it has been observed from molecular CC calculations that the error in energy decreases asymptotically as X^{-3}, while the cardinality of the basis set increases as X^3 (cf. [24, 26]). This observation is consistent with the convergence rate derived from partial wave expansions for atoms.

In order to compare the computational complexity with best N-term approximations, we consider the CC method restricted to *single and double excitations* (CCSD) (cf. [25]). This is roughly equivalent to a Jastrow factor (1.2) where only one- and two-particle correlation functions are taken into account. For CCSD calculations the number of degrees of freedom N, represented by the cluster amplitudes, increases as X^6, which means that the corresponding error in energy decreases as $N^{-1/2}$. Let us assume that our actual realization of the Jastrow factor ansatz has the same convergence behaviour as a strictly variational treatment according to Rayleigh-Ritz's variational principle. Such kind of assumption is actually fulfilled in practice for CC and related methods like many-body perturbation theory. According to a standard argument, the error in energy has twice the convergence rate with respect to the H^1 error. Therefore best N-term approximation converges as N^{-1} with respect to the energy compared with $N^{-1/2}$ for standard GTO basis sets.

We want to close this section with a few remarks on how to overcome these limitations for both GTO and wavelet bases. A promising approach for GTO basis sets is to incorporate a term linear in the inter-electron distance into the wavefunction (cf. [35]). This yields twice the convergence rate and enables very accurate calculations using the CC method or many-body perturbation theory (cf. [24, 26]). For wavelets we can use a similar approach which, however, offers a larger flexibility due to the peculiar features of local multi-resolution bases. In order to illustrate the difference between wavelet and GTO basis sets it is sufficient to consider a single atom. Increasing the size of an atomic GTO basis set always corresponds to a global refinement in the configuration space of electrons. Compared to this, wavelets, on sufficiently fine scales, allow for local refinements in the immediate neighbourhood of the $e - e$ cusp. Concerning two-particle correlation functions, the $e - e$ cusp corresponds to a three dimensional hypersurface in $\mathbb{R}^6$, and therefore requires a comparatively large number of tensor-product wavelets (2.4) on fine scales. However it is well known from QMC methods that Jastrow factors actually vary rather smoothly along the $e - e$ cusp (cf. [15, 16, 22]). Therefore it is possible to considerably reduce the number of degrees of freedom by contracting wavelets in the immediate neighbourhood of the $e - e$ cusp [12, 37]. Such kind of contraction scheme can be represented through a hierarchical Kronecker tensor product decomposition (cf. [23]), which, furthermore, enables an efficient evaluation of integrals (cf. [37]).

References

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

[2] C.E. Campbell, E. Krotscheck, and T. Pang, *Electron correlations in atomic systems*, Physics Reports **223** (1992) 1-42.

[3] D. Ceperley, *Ground state of the fermion one-component plasma: A Monte Carlo study in two and three dimensions*, Phys. Rev. B **18** (1978) 3126-3138.

[4] J.W. Clark, *Variational theory of nuclear matter*, in Progress in Nuclear and Particle Physics Vol. 2, edited by D. H. Wilkinson.Pergamon, Oxford, 1979, pp. 89-199.

[5] E.T. Copson, *Asymptotic Expansions*. Cambridge University Press, Cambridge, 1967.

[6] W. Dahmen, S. Prößdorf, and R. Schneider, *Wavelet approximation methods for pseudodifferential equations I: Stability and convergence* Math. Z. **215**, (1994) 583-620 (1994).

[7] W. Dahmen, S. Prößdorf, and R. Schneider, *Wavelet approximation methods for pseudodifferential equations II: Matrix compression and fast solution*, Adv. Comp. Maths. **1** (1993) 259-335.

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

[9] R.A. DeVore, B. Jawerth, and V. Popov, *Compression of wavelet decompositions*, Amer. J. Math. **114** (1992) 737-785.

[10] R.A. DeVore, S.V. Konyagin, and V.N. Temlyakov, *Hyperbolic wavelet approximation*, Constr. Approx. **14** (1998) 1-26.

[11] N.D. Drummond, M.D. Towler, and R.J. Needs, *Jastrow correlation factor for atoms, molecules, and solids* Phys. Rev. B **70** (2004), 235119, 11 pages.

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

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

[14] H.-J. Flad, W. Hackbusch, and R. Schneider, *Best N-term approximation in electronic structure calculations. I. One-electron reduced density matrix*, Preprint 60, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig, 2005.

[15] H.-J. Flad and A. Savin, *Transfer of electron correlation from the electron gas to inhomogeneous systems via Jastrow factors*, Phys. Rev. A. **50** (1994) 3742-3746.

[16] H.-J. Flad and A. Savin, *A new Jastrow factor for atoms and molecules, using two-electron systems as a guiding principle*, J. Chem. Phys. **103** (1995) 691-697. (1995) 691-697.

[17] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, and T. Ostergaard Sorensen, *Sharp regularity results for Coulombic many-electron wave functions*, Commun. Math. Phys. **255** (2005) 183-227.

[18] D.E. Freund, B.D. Huxtable, and J.D. Morgan III, *Variational calculations on the helium isoelectronic sequence*, Phys. Rev. A **29** (1984) 980-982.

[19] P. Fulde, *Electron Correlations in Molecules and Solids*, 2nd edition. Springer, Berlin, 1993.

[20] P. Fulde, *Ground-state wave functions and energies of solids*, Int. J. Quant. Chem. **76** (2000) 385-395.

[21] J. Garcke and M. Griebel, *On the computation of the eigenproblems of hydrogen and helium in strong magnetic and electric fields with the sparse grid combination technique*, J. Comp. Phys. **165** (2000) 694-716.

[22] R. Gaudoin, M. Nekovee, W.M.C. Foulkes, R.J. Needs, and G. Rajagopal, *Inhomogeneous random-phase approximation and many-electron trial wave functions*, Phys. Rev. B **63** (2001), 115115, 15 pages.

[23] W. Hackbusch, B.N. Khoromskij, and E. Tyrtyshnikov, *Hierarchical Kronecker tensor-product approximation*, Preprint 35, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig 2003.

[24] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A.K. Wilson, *Basis-set convergence in correlated calculations on Ne, N_2, and H_2O*, Chem. Phys. Lett. **286** (1998) 243-252.

[25] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory*. Wiley, New York, 1999.

[26] T. Helgaker, W. Klopper, H. Koch, and J. Noga, *Basis-set convergence of correlated calculations on water*, J. Chem. Phys. **106**, (1997) 9639-9646.

[27] R.N. Hill, *Rates of convergence and error estimation formulas for the Rayleigh-Ritz variational method*, J. Chem. Phys. **83** (1985) 1173-1196.

[28] M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, and H. Stremnitzer, *Local properties of Coulombic wave functions*, Commun. Math. Phys. **163** (1994) 185-215.

[29] M. Hoffmann-Ostenhof and R. Seiler, *Cusp conditions for eigenfunctions of n-electron systems*, Phys. Rev. A **23** (1981) 21-23.

[30] C.-J. Huang, C.J. Umrigar, and M.P. Nightingale, *Accuracy of electronic wave functions in quantum Monte Carlo: The effect of high-order correlations*, J. Chem. Phys. **107** (1997) 3007-3013.

[31] T. Kato, *On the eigenfunctions of many-particle systems in quantum mechanics*, Commun. Pure Appl. Math. **10** (1957) 151-177.

[32] E. Krotscheck, *Variations on the electron gas*, Ann. Phys. (N.Y.) **155** (1984) 1-55.

[33] E. Krotscheck, *Theory of inhomogeneous quantum systems. III. Variational wave functions for Fermi fluids*, Phys. Rev. B **31** (1985) 4267-4278.

[34] E. Krotscheck, W. Kohn, and G.-X. Qian, *Theory of inhomogeneous quantum systems. IV. Variational calculations of metal surfaces*, Phys. Rev. B **32** (1985) 5693-5712.

[35] W. Kutzelnigg, *r_{12}-Dependent terms in the wave function as closed sums of partial wave amplitudes for large l*, Theoret. Chim. Acta **68**, (1985) 445-469.

[36] W. Kutzelnigg and J.D. Morgan III, *Rates of convergence of the partial-wave expansions of atomic correlation energies*, J. Chem. Phys. **96** (1992) 4484-4508.

[37] H. Luo, D. Kolb, H.-J. Flad, W. Hackbusch, and T. Koprucki, *Wavelet approximation of correlated wavefunctions. II. Hyperbolic wavelets and adaptive approximation schemes*, J. Chem. Phys. **117** (2002) 3625-3638.

[38] H. Luo, D. Kolb, H.-J. Flad, and W. Hackbusch, *Perturbative calculation of Jastrow factor*, in preparation.

[39] P.-A. Nitsche, *Sparse approximation of singularity functions*, Constr. Approx. **21** (2005) 63-81.

[40] P.-A. Nitsche, *Best N term approximation spaces for sparse grids*, Research Report No. 2003-11, Seminar für Angewandte Mathematik, ETH Zürich.

[41] T. Pang, C.E. Campbell, and E. Krotscheck, *Local structure of electron correlations in atomic systems*, Chem. Phys. Lett. **163** (1989) 537-541.

[42] K.E. Schmidt and J.W. Moskowitz, *Correlated Monte Carlo wave functions for the atoms He through Ne*, J. Chem. Phys. **93** (1990) 4172-4178.

[43] E.M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals.* Princeton University Press, 1993.

[44] G. Stollhoff, *The local ansatz extended*, J. Chem. Phys. **105** (1996) 227-234.

[45] G. Stollhoff and P. Fulde, *On the computation of electronic correlation energies within the local approach* J. Chem. Phys. **73** (1980) 4548-4561.

[46] J.D. Talman, *Linked-cluster expansion for Jastrow-type wave functions and its application to the electron-gas problem*, Phys. Rev. A, **10** (1974) 1333-1344.

[47] J.D. Talman, *Variational calculation for the electron gas at intermediate densities*, Phys. Rev. A, **13** (1976) 1200-1208.

[48] C.J. Umrigar, K.G. Wilson, and J.W. Wilkins, *Optimized trial wave functions for quantum Monte Carlo calculations*, Phys. Rev. Lett. **60** (1988) 1719-1722.

[49] A.J. Williamson, S.D. Kenny, G. Rajagopal, A.J. James, R.J. Needs, L.M. Fraser, W.M.C. Foulkes, and P. Maccallum, *Optimized wavefunctions for quantum Monte Carlo studies of atoms and solids*, Phys. Rev. B **53** (1996) 9640-9648.

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

[51] H. Yserentant, *Sparse grid spaces for the numerical solution of the electronic Schrödinger equation*, Numer. Math. to be published.

Appendix

A Multi-resolution analysis within the random phase approximation

Among numerous applications of Jastrow factors in quantum chemistry and condensed matter physics, the homogenous electron gas is probably the most extensively studied model system. It can be considered as a simplified model for metals (cf. [19]) where individual atomic cores are replaced by a constant background charge density. Due to translational invariance and isotropy of this model system it became possible to study Jastrow factors using various kinds of quantum many-body theories (cf. [3, 32, 46, 47]). In the following we want to consider Jastrow factors within the so-called *random phase approximation* (RPA). For a thorough discussion of the underlying physics, we refer, e.g., to [19]. This approximation provides an explicit representation of the two-particle correlation function in momentum space,

$$\hat{\mathcal{F}}_{\mathrm{RPA}}(k) = \frac{1}{2\rho S_F(k)} \left[1 - \left(1 + \frac{4\rho\hat{v}(k)S_F(k)^2}{k^2} \right)^{\frac{1}{2}} \right] \tag{A.1}$$

(cf. [46]), where the static structure function of a non-interacting particle system

$$S_F(k) = \begin{cases} \frac{3}{4}\frac{k}{k_F} - \frac{1}{16}\left(\frac{k}{k_F}\right)^3 & k < 2k_F, \\ 1 & k \geq 2k_F, \end{cases} \tag{A.2}$$

enters as well as the Fourier transform of the Coulomb potential $\hat{v} = 4\pi/k^2$. For the sake of notational simplicity, we have introduced the variable $k := |\mathbf{k}|$ to denote the absolute value of the relative momentum $\mathbf{k} \in \mathbb{R}^3$ between two electrons. The homogeneous electron gas can be either characterized by its density ρ or by the Fermi momentum k_F, both quantities are related through $\rho = k_F^3/(3\pi^2)$.

In order to study the asymptotic behaviour of the RPA two-particle correlation function in configuration space, we perform a Littlewood-Paley decomposition of (A.1) in momentum space. Starting with a cut-off function $\eta(k) \in C_0^\infty(\overline{\mathbb{R}}_+)$ with $\eta(k) = 1$ for $k < k_F/2$ and $\eta(k) = 0$ for $k \geq k_F$, we construct a dyadic partition of unity in momentum space (cf. [43]):

$$1 = \eta(k) + \sum_{j=1}^\infty \delta_j(k) \qquad \text{with} \quad \delta_j(k) := \eta(2^{-j}k) - \eta(2^{-j+1}k).$$

Inserting this partition of unity into Eq. (A.1), we obtain a multi-scale decomposition of the two-particle correlation function

$$\mathcal{F}_{\mathrm{RPA}}(r) = S_0\mathcal{F}(r) + \sum_{j=0}^\infty \Delta_j\mathcal{F}(r), \tag{A.3}$$

with respect to its energy scales, where individual terms are given via the inverse Fourier transforms

$$S_0\mathcal{F}(r) := \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} e^{i\mathbf{k}\mathbf{r}}\eta(k)\hat{\mathcal{F}}_{\mathrm{RPA}}(k)\,d\mathbf{k}, \tag{A.4}$$

$$\Delta_j\mathcal{F}(r) := \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} e^{i\mathbf{k}\mathbf{r}}\delta_j(k)\hat{\mathcal{F}}_{\mathrm{RPA}}(k)\,d\mathbf{k}.$$

To simplify our notation, we have introduced the variable $r := |\mathbf{r}|$ to denote the absolute value of the vector $\mathbf{r} \in \mathbb{R}^3$ describing the relative position of two electrons in configuration space. First we want to discuss the behaviour of the high-energy scales in configuration space. For this we consider the partial sum

$$\mathcal{F}_s(r) := \sum_{j=j_0}^\infty \Delta_j\mathcal{F}(r), \tag{A.5}$$

where j_0 is chosen in such a way that $\min \mathrm{supp}\,\delta_{j_0} > 2k_F$. The sum (A.5) essentially comprises short-range correlations close to the $e - e$ cusp.

Proposition A.1 *The short-range part of the RPA two-particle correlation function $\mathcal{F}_s(|\mathbf{x} - \mathbf{y}|)$ satisfies Assumption 1.1. Furthermore it satisfies the generalized asymptotic smoothness property*

$$\left| \partial_{\mathbf{x}}^\alpha \partial_{\mathbf{y}}^\beta \mathcal{F}_s(|\mathbf{x} - \mathbf{y}|) \right| \leq C_{\alpha,\beta} |\mathbf{x} - \mathbf{y}|^{1-|\alpha|-|\beta|-N}, \qquad \textit{if } \mathbf{x} \neq \mathbf{y} \textit{ and } |\alpha| + |\beta| + N \geq 1, \tag{A.6}$$

for arbitrary $N \geq 0$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$.

Proof. According to our construction, we first observe for the Fourier transform $\hat{\mathcal{F}}_s(k) \in C^\infty(\mathbb{R}_+)$ with $\hat{\mathcal{F}}_s(k) = 0$ for $k \leq 2k_F$. In addition, it satisfies the estimate

$$\left| \partial_{\mathbf{k}}^\alpha \hat{\mathcal{F}}_s(k) \right| \leq C_\alpha (1+k)^{-4-|\alpha|} \qquad \text{for } |\alpha| \geq 0,$$

as can be easily seen from Eq. (A.1). Therefore $\hat{\mathcal{F}}_s$ can be considered as a translationally invariant symbol of a pseudo-differential operator in the symbol class $S^{-4}(\mathbb{R}^3 \times \mathbb{R}^3)$ following Stein's [43] approach to pseudo-differential operators in $\mathbb{R}^n$. The estimate (A.6) for $|\alpha| + |\beta| + N > 1$ is an immediate consequence of [43, Proposition VI.4.1, p. 241]. It remains to prove the boundedness of the first derivatives. For this we consider first the two-particle correlation function from Corollary 2.4

$$\mathcal{F}_c^{(2)}(|\mathbf{x} - \mathbf{y}|) = \frac{1}{2}|\mathbf{x} - \mathbf{y}|e^{-|\mathbf{x}-\mathbf{y}|^2}, \tag{A.7}$$

where an appropriate cut-off function has been added. The prefactor corresponds to Kato's cusp condition for two electrons with antiparallel spin [31]. Performing the Fourier transformation, we can apply formula (3.5) and obtain the asymptotic expansion

$$\hat{\mathcal{F}}_c^{(2)}(k) = \int_{\mathbb{R}^3} e^{-i\mathbf{k}\mathbf{r}} \mathcal{F}_c^{(2)}(|\mathbf{r}|)\,d\mathbf{r} = \frac{2\pi}{k} \int_0^\infty s^2 e^{-s^2} \sin(ks)\,ds \qquad \text{with } k := |\mathbf{k}|$$

$$\sim -\frac{4\pi}{k^4} - \frac{48\pi}{k^6} + O\left(\frac{1}{k^8}\right) \qquad \text{for } k \to \infty. \tag{A.8}$$

This has to be compared with the asymptotic expansion of the RPA two-particle correlation function (A.1)

$$\hat{\mathcal{F}}_{\text{RPA}}(k) = \frac{1}{2\rho}\left(1 - \sqrt{1 + \frac{16\pi\rho}{k^4}}\right) \qquad \text{for } k > 2k_F$$

$$\sim -\frac{4\pi}{k^4} + O\left(\frac{1}{k^8}\right). \tag{A.9}$$

We can decompose the short-range part (A.5) according to

$$\mathcal{F}_s = \mathcal{F}_c + \underbrace{\mathcal{F}_s - \mathcal{F}_c}_{:=\mathcal{F}_{s-c}}.$$

It follows from the asymptotic expansions (A.8) and (A.9) that $\hat{\mathcal{F}}_{s-c} \sim O\left(\frac{1}{k^6}\right)$. Therefore $\mathcal{F}_{s-c}(|\mathbf{x} - \mathbf{y}|) \in C^1(\mathbb{R}^3 \times \mathbb{R}^3)$, which demonstrates together with Eq. (A.7) the boundedness of the first derivatives. ∎

At next we consider the low-energy contribution $S_0\mathcal{F}$ which gives rise to long-range correlations in configuration space. Such kind of correlations are typical for metals (cf. [19]) and seem to be less important or even absent in semiconductors and insulators (cf. [20]). It is an immediate consequence of definition (A.4) that $S_0\mathcal{F}(|\mathbf{x} - \mathbf{y}|) \in C^\infty(\mathbb{R}^3 \times \mathbb{R}^3)$. Furthermore, the long-range part $S_0\mathcal{F}$ satisfies an asymptotic smoothness property which allows to estimate wavelet coefficients of the two-particle correlation function at large distances.

Proposition A.2 *The long-range part of the RPA correlation function $S_0\mathcal{F}$, defined via the Littlewood-Paley decomposition (A.3), satisfies the estimate*

$$\left| \partial_{\mathbf{x}}^\alpha \partial_{\mathbf{y}}^\beta S_0\mathcal{F}(|\mathbf{x} - \mathbf{y}|) \right| \leq C_{\alpha,\beta} (1 + |\mathbf{x} - \mathbf{y}|)^{-(1+|\alpha|+|\beta|)}. \tag{A.10}$$

Proof. The proof is based on an analysis of the asymptotic behaviour of $S_0\mathcal{F}(r)$ and its derivatives for $r \to \infty$. Using spherical symmetry, we can express the Fourier integral (A.4) in the form

$$S_0\mathcal{F}(r) = \frac{1}{2\pi^2 r} \int_0^\infty k^{-1}\eta(k)\left(k^2\hat{\mathcal{F}}(k)\right)\sin(kr)\,dk. \qquad (A.11)$$

From this we can derive a recurrence relation for the derivatives with respect to r,

$$\frac{d^p S_0\mathcal{F}(r)}{dr^p} = -\frac{p}{r}\frac{d^{p-1}S_0\mathcal{F}(r)}{dr^{p-1}} + \frac{1}{2\pi^2 r}\int_0^\infty k^{-1}\eta(k)\left(k^2\hat{\mathcal{F}}(k)\right)\frac{\partial^p \sin(kr)}{\partial r^p}\,dk. \qquad (A.12)$$

It is obvious from Eq. (A.1) that the function $k^2\hat{\mathcal{F}}(k)\eta(k)$ can be extended to a function in $C_0^\infty(\mathbb{R})$. With this it is possible to use the standard formula (3.5) for asymptotic expansions of oscillatory integrals. The expansion is immediately applicable to the second term on the right hand side of the recurrence relation (A.12). The integral (A.11) requires the case $n = -1$ for which the asymptotic expansion (3.5) does not apply. However it is straightforward to modify the proof of the asymptotic expansion (3.5) for the integral (A.4) (see e.g. [43, Paragraph VIII.5.1]) because it contains only $\sin(kr)$ which is of $O(k)$ for $k \to 0$. The modified asymptotic expansion for $r \to \infty$,

$$S_0\mathcal{F}(r) \sim \frac{\tilde{a}_0}{r} + \sum_{j=0}^\infty \frac{\tilde{a}_{2j+1}}{r^{2j+2}} \qquad \text{with } \tilde{a}_0 = \frac{1}{\sqrt{4\pi\rho}}, \quad \tilde{a}_j = \frac{i^{j-1}}{2\pi^2 j}\left.\frac{d^j\left(k^2\hat{\mathcal{F}}(k)\right)}{dk^j}\right|_{k=0} \qquad (j > 0).$$

contains in leading order the plasmon frequency of the homogeneous electron gas $\omega_{\mathrm{pl}} = \sqrt{4\pi\rho}$ (cf. [19]). Altogether we derived asymptotic expansions for the long-range part of the RPA correlation function and its derivatives via the recurrence relation (A.12), which proves the asymptotic smoothness property (A.10). ∎

An important consequence of Proposition A.2 is that there exists a multi-resolution representation for the long-range part of the RPA correlation function (A.4) that can be further compressed without significant loss of accuracy. Using techniques similar to those used in Lemma 2.1, one of us proved the estimate

$$\left|\iint \gamma_{j,\mathbf{a}}^{(\mathbf{s})}(\mathbf{x})S_0\mathcal{F}(|\mathbf{x}-\mathbf{y}|)\gamma_{\ell,\mathbf{b}}^{(\mathbf{t})}(\mathbf{y})\,d\mathbf{x}d\mathbf{y}\right| \lesssim \frac{2^{-j(3/2+p|\mathbf{s}|)}2^{-\ell(3/2+p|\mathbf{t}|)}}{\mathrm{dist}(\mathrm{supp}\,\gamma_{j,\mathbf{a}}^{(\mathbf{s})}, \mathrm{supp}\,\gamma_{\ell,\mathbf{b}}^{(\mathbf{t})})^{1+p(|\mathbf{s}|+|\mathbf{t}|)}},$$

in the context of Galerkin discretization of pseudo-differential operators (cf. [6, 7]).

Some of the remaining terms $\Delta_j\mathcal{F}$ ($j = 1, \ldots, j_0 - 1$) of the Littlewood-Paley decomposition (A.3), with $2k_F \in \mathrm{supp}\,\delta_j$, have discontinuous second derivatives at $k = 2k_F$. This is an immediate consequence from the corresponding behaviour of the static structure function of a non-interacting particle system (A.2) that enters into Eq. (A.1). Concerning their asymptotic behaviour in configuration space we get

$$\sum_{j=1}^{j_0-1}\Delta_j\mathcal{F}(r) \sim o\left(\frac{1}{r^2}\right),$$

which means that these correlations decay much faster than the long-range correlations discussed above. Nevertheless, the lack of an asymptotic smoothness property similar to the estimate (A.10) indicates the presence of an oscillatory behaviour in configuration space which prevents an efficient compression of the remaining terms in a multi-resolution basis. Almost nothing is known about their physical significance. Comparative QMC calculations between RPA and parametrized Jastrow factors that do not possess any oscillatory behaviour indicate however that these correlations are of minor importance (cf. [3]).