ESAIM: M2AN
Vol. 41, N° 2, 2007, pp. 215–247
DOI: 10.1051/m2an:2007015

**ESAIM: Mathematical Modelling and Numerical Analysis**
www.edpsciences.org/m2an

# SPARSE GRIDS FOR THE SCHRÖDINGER EQUATION

Michael Griebel[1] and Jan Hamaekers[1]

**Abstract.** We present a sparse grid/hyperbolic cross discretization for many-particle problems. It involves the tensor product of a one-particle multilevel basis. Subsequent truncation of the associated series expansion then results in a sparse grid discretization. Here, depending on the norms involved, different variants of sparse grid techniques for many-particle spaces can be derived that, in the best case, result in complexities and error estimates which are independent of the number of particles. Furthermore we introduce an additional constraint which gives antisymmetric sparse grids which are suited to fermionic systems. We apply the antisymmetric sparse grid discretization to the electronic Schrödinger equation and compare costs, accuracy, convergence rates and scalability with respect to the number of electrons present in the system.

**Mathematics Subject Classification.** 35J10, 65N25, 65N30, 65T40, 65Z05.

Received December 13, 2005.

## Introduction

In this article we consider the electronic Schrödinger equation (first without spin for reasons of simplicity)

$$H\psi(\mathbf{x}_1,\ldots,\mathbf{x}_N) = E\psi(\mathbf{x}_1,\ldots,\mathbf{x}_N) \tag{1}$$

with the Hamilton operator

$$H = T + V \quad \text{where} \quad T = -\frac{1}{2}\sum_{i=1}^{N}\Delta_i$$

and

$$V = \begin{cases} +\displaystyle\sum_{i=1}^{N}\sum_{j=1}^{N_{nuc}} Z_j|\mathbf{x}_i-\mathbf{R}_j|_2 - \sum_{i=1}^{N}\sum_{j>i}^{N}|\mathbf{x}_i-\mathbf{x}_j|_2, & d=1,\\ +\displaystyle\sum_{i=1}^{N}\sum_{j=1}^{N_{nuc}} Z_j\log(|\mathbf{x}_i-\mathbf{R}_j|_2) - \sum_{i=1}^{N}\sum_{j>i}^{N}\log|\mathbf{x}_i-\mathbf{x}_j|_2, & d=2,\\ -\displaystyle\sum_{i=1}^{N}\sum_{j=1}^{N_{nuc}} \frac{Z_j}{|\mathbf{x}_i-\mathbf{R}_j|_2} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{|\mathbf{x}_i-\mathbf{x}_j|_2}, & d=3. \end{cases}$$

*Keywords and phrases.* Schrödinger equation, numerical approximation, sparse grid method, antisymmetric sparse grids.

[1] Institute of Numerical Simulation, University of Bonn, Wegelerstraße 6, 53115 Bonn, Germany. griebel@ins.uni-bonn.de; hamaekers@ins.uni-bonn.de

Here, $\mathbf{x}_i := (x_{1,i}, \ldots, x_{d,i}) \in \mathbb{R}^d$ denotes the position of the $i$-th electron, $i = 1, \ldots, N$, and $\mathbf{R}_j \in \mathbb{R}^d$ denotes the fixed position of the $j$-th nucleus, $j = 1, \ldots, N_{nuc}$. The operator $\Delta_i$ is the Laplacian acting on the $\mathbf{x}_i$-component of $\psi$, *i.e.* $\Delta_i = \sum_{j=1}^{d} \partial^2/\partial(x_{j,i})^2$, $Z_j$ is the charge of the $j$-th nucleus and the norm $|.|_2$ denotes the usual Euclidean distance in $\mathbb{R}^d$. The solution $\psi$ describes the wave function associated to the eigenvalue $E$. Note that for $d = 1$ and $d = 2$ we use the associated Coulomb interaction potential with reversed sign to account for a repulsion of electrons and for an attraction of nuclei and electrons.

This eigenvalue problem results from the Born-Oppenheimer approximation [65] to the general Schrödinger equation for a system of electrons and nuclei which takes the different masses of electrons and nuclei into account. It is one of the core problems of computational chemistry. Its successful treatment would allow to predict the properties of arbitrary atomic systems and molecules [26]. However, except for very simple cases, there is no analytical solution for (1) available. Also a direct numerical approach is impossible since $\psi$ is a $d \cdot N$-dimensional function. Any discretization on *e.g.* uniform grids with $O(K)$ points in each direction would involve $O(K^{d \cdot N})$ degrees of freedoms which are impossible to store for $d = 3, N > 1$. Furthermore, only a convergence rate of the type

$$\|\psi - \psi_K\|_{\mathcal{H}^s} \leq c(N, d) \cdot K^{-r/(d \cdot N)} \|\psi\|_{\mathcal{H}^{s+r}}$$

can be achieved, where $\|.\|_{\mathcal{H}^s}$ is the usual Sobolev norm in $\mathcal{H}^s$, $r$ denotes the isotropic smoothness of $\psi$ and $c$ is a constant which may depend on $N$ and $d$ but not on $K$. Here, we encounter the curse of dimensionality [7], *i.e.* the rate of convergence deteriorates exponentially with the dimension $d$ and the number $N$ of electrons, respectively.

Therefore, most approaches resort to an approximation of (1) only. Examples are the classical Hartree-Fock method or its successive refinements like configuration interaction or coupled clusters which lead to tractable approximations to Schrödinger's equation. For the conventional hierarchy of additive methods, finite sums of Slater determinants are used. Some existence and convergence theory can be found in [33,57,58]. Density functional theory provides an alternative framework. Here, the high dimensionality is traded for a highly nonlinear equation on the one-particle density of the ground state, with an unknown but principally exact exchange-correlation part. There, the Kohn-Sham equations allow to express the kinetic energy in terms of the density. This highly successful method is widely employed but cannot be improved upon systematically. Furthermore there are the reduced density matrix (RDM) [64] and the r12 approach [29] which lead to improved accuracy and open the way to new applications. A survey of these methods can be found in [4,59,60]. A major problem with these techniques is that, albeit quite successful in practice, they nevertheless only provide approximations. Also a systematical improvement is usually not easily available such that convergence of the model to Schrödinger's equation is achieved.

In this article, we intend to directly discretize the Schrödinger equation without resorting to any model approximation. To this end, we aim at a discretization method which circumvents the above-mentioned curse of dimensionality. We envision a convergence rate of the type

$$\|\psi - \psi_K\|_{\mathcal{H}^s} \leq c(N, d) \cdot K^{-r/d} \|\psi\|_{\mathcal{H}^{s+r}_{\text{mix}}} \tag{2}$$

where the rate of convergence does no longer exponentially deteriorate with the number $N$ of particles. Now, however, a more restrictive smoothness requirement, namely the boundedness of a certain $(s + r)$-th mixed derivative may be involved. Such favourable convergence properties can be achieved for various sparse grid discretization methods in the context of integration problems [35, 36], integral equations [32, 43] and elliptic partial differential equations, see [12] and the references cited therein. In Fourier space, such methods are also known under the name hyperbolic cross approximation.

The derivation of a sparse grid method starts from a one-dimensional multiscale basis which exhibits an appropriate decay property for the coefficients of a function representation. A tensor product construction then results in a multilevel basis for the multivariate case. Subsequent truncation of the corresponding multivariate series expansion finally leads to a sparse grid discretization scheme. Here, depending of the norm and truncation strategy used, different variants of sparse grids (regular sparse grids, energy-norm based sparse grids, dimension

adaptive sparse grids, see [12]) can be derived which, in the best case, result in convergence rates of the type (2) provided that certain mixed derivatives of the solution are bounded.

Recently Yserentant showed in [77] that such a type of smoothness prerequisite (involving $\mathcal{H}^{1,1}_{\text{mix}}$, $\mathcal{H}^{\frac{1}{2},1}_{\text{mix}}$, see (11)), is indeed valid for the solution $\psi$ of Schrödinger's equation. This result suggests that it is possible to numerically solve (1) up to a prescribed accuracy with an amount of work which does not scale exponentially in the number $N$ of electrons with respect to $K$. Note that in the case $d = 3$ the Coulomb potential $\frac{1}{|\mathbf{x}-\mathbf{y}|_2}$ is unbounded at the set of coalescence points

$$\left\{ \vec{\mathbf{x}} \in (\mathbb{R}^d)^N \; : \; \left( \prod_{i=1}^{N} \prod_{j=1}^{N_{nuc}} |\mathbf{x}_i - \mathbf{R}_j|_2 \right) \left( \prod_{i=1}^{N} \prod_{j>i}^{N} |\mathbf{x}_i - \mathbf{x}_j|_2 \right) = 0 \right\}$$

and thus the eigenfunctions are nonanalytic there. In 1957 Kato studied the solution near two-particle coalescence points and showed that the eigenfunctions are locally Lipschitz [52]. Further results on the regularity of the eigenfunctions of the Schrödinger operator and on the behaviour of a many-electron wavefunction in the neighborhood of the coalescence points (*e.g.* cusp conditions) can be found in [29–31, 50, 52].

In this article we develop and study a generalized sparse grid/hyperbolic cross technique for the electronic Schrödinger equation (1). For reasons of simplicity we restrict ourselves to the setting of a $d \cdot N$-dimensional product domain $\Omega = I^{d \cdot N}$ with $I = [0, 2\pi]$ and periodic boundary conditions. As a multilevel basis for the one-particle space we use the Fourier function system $\{(\frac{1}{2\pi})^{Nd/2} \mathrm{e}^{i\mathbf{k}^T\mathbf{x}}, \mathbf{k} \in \mathbb{Z}^d, \mathbf{x} \in I^d\}$. Beside its well known decay properties for sufficiently smooth functions, this choice provides an orthogonal basis of $\mathcal{L}^2(\Omega)$. This is advantageous in the Galerkin discretization process since, due to the Slater-Condon rules, we then obtain a sparse matrix for the discretized eigenvalue problem. Note that our approach is by no means restricted to this specific choice of multilevel basis and to the periodic setting. Any multilevel basis for the one-particle space with a sufficient decay property may be used as basic ingredient for our sparse grid approach with similar results. Candidates are other hierarchical global polynomial systems [8, 10, 12, 51, 73] or function families with localization properties like wavelets [18], prewavelets [15, 42], interpolets [19, 22], and related wavelet-like constructs, see [12, 16] for a survey. But also multiscale finite element systems and frames [37, 40, 41, 67] or multiscale Gaussians [63] may be used[1,2].

From a basis for the one-particle space we derive various sparse grid/hyperbolic cross spaces for the $N$-particle space by means of a tensor product construction and a subsequent series truncation. Here we first consider functions from spaces with bounded mixed derivatives $\mathcal{H}^t_{\text{mix}}$ and construct regular sparse grids for $N$-particle spaces. We derive estimates for the dimension of the associated discrete space $V_{K,0}$ and the associated error. The degrees of freedoms then scale logarithmically with the number $N$ of particles with respect to the discretization parameter $K$. We then introduce an additional parameter $T$ which allows us to generalize the sparse grid approach. It makes it possible to switch to discrete spaces $V_{K,T}$ which can be chosen optimally for functions from spaces with bounded mixed derivatives $\mathcal{H}^{t,l}_{\text{mix}}$ where the parameter $t$ relates to mixed derivatives and $l$ relates to the (isotropic) partial derivatives of degree $l$. We derive estimates for the dimension of $V_{K,T}$ and the associated approximation error. In special cases, *i.e.* if $T \in (0, 1]$ the dependence of the dimension of $V_{K,T}$ on $N$ with respect to the asymptotics in $K$ is completely removed while the order of approximation is for $T \leq (s - l)/t$ the same as in the full grid case. Here the approximation error is measured in the classical Sobolev norm $\|.\|_{\mathcal{H}^s}$.

We then restrict our sparse grid approach to the case of antisymmetric functions which obey Pauli's principle. To this end, we replace the conventional (inner) tensor product by an outer product which involves Slater

[1]Note that a wavelet-like system with localization properties might even further improve the complexity (Besov spaces, see [66]) when it comes to the adaptive local resolution of nuclei-electron cusps and electron-electron cusps, compare also [27, 28].

[2]Also, instead of a product approach which results in $d$-dimensional functions for the one-particle space with associated multivariate index $\mathbf{k}$ and *e.g.* anisotropic local supports, we can use isotropic constructions like in conventional $d$-dimensional hierarchical bases or conventional isotropic wavelets which possess only a univariate index for the notation of the level (but involve $2^d - 1$ mother scaling functions).

determinants. Then, many degrees of freedom of the conventional sparse grid construction coincide, *i.e.* their associated basis functions are now (up to sign) the same. The idea is to only take one representative of the indices with coinciding basis functions into consideration. This leads to an additional constraint on the indices of the sparse grid basis functions which results in a further substantial reduction of the complexity, *i.e.* the number of degrees of freedom is basically reduced by the factor $N!(N-S)!$ where $S$ denotes the number of electrons with negative spin and $N-S$ is the number of electrons with positive spin. The estimate for the associated error is the same as in the non-constrained case.

We use the resulting basis functions of the antisymmetric sparse grid space in a Galerkin discretization of (1). Due to the orthogonality a large number of the entries of the associated stiffness matrix is zero which reduces the storage requirements and complexity substantially. Furthermore, since we use a Fourier basis, the inner products, *i.e.* the non-zero entries of the matrix, can be reduced to $d$- and $2d$-dimensional integrals, whose values can be computed either numerically or, in the case $d = 1$ and $d = 3$ even analytically. We then solve the discrete eigensystem by a parallel Lanczos method.

We finally apply our new antisymmetric sparse grid approach to model problems with varying number of electrons, and compare costs, accuracy, convergence rates and scalability with respect to the number of electrons present in the system. Altogether, we give a direct method for Schrödinger's equation without resorting to any model approximation. The errors involved in our approach are pure discretization errors for which a provable convergence rate is provided.

The remainder of this paper is organized as follows: in Section 1 we discuss the conventional sparse grid approach, here for the case of one-particle spaces as basic building blocks. We employ for the $d$-dimensional one-particle space on $I^d$ the anisotropic product of a 1D multiscale basis on $I$. As an example we use here for reasons of simplicity the Fourier basis. For a $N$-particle system a further product approach then leads to a multiscale basis on $I^{d\cdot N}$. Truncation leads to different variants of sparse grid subspaces. Here, besides the conventional sparse grid approach we focus on optimized sparse grids which allow to take advantage of certain mixed smoothness properties of the function to be represented. We discuss the associated complexities and approximation properties. In Section 2 we generalize the sparse grid approach to the case of antisymmetry. To this end, the conventional product is first replaced by the outer product which involves the Slater determinant construction. Then, we impose additional conditions on the level indices of the multivariate basis which reflect the Pauli principle. We thus obtain a true basis for antisymmetric sparse grid spaces with a substantially reduced amount of degree of freedoms. For different variants of such antisymmetric sparse grid spaces we derive the associated complexities and approximation properties. Then, in Section 4 we apply the Galerkin approach for the electronic Schrödinger equation using antisymmetric sparse grid spaces. We set up the stiffness matrix, consider its non-zero structure which results from the Slater-Condon rules and solve the associated discrete eigenvalue problem with a Lanczos solver. In Section 5 we apply our approach to model problems with varying number of electrons and compare costs, accuracy, convergence rate and scalability with respect to the number of electrons present in the system. Finally we give some concluding remarks in Section 6.

## 1. Sparse grids for particle spaces

In the following we introduce various Sobolev spaces and norms for particles. We then introduce approximation spaces related to regular sparse grids and discuss their dimension and approximation rates. Finally we consider optimized sparse grid spaces and derive their dimension and approximation rates. For reasons of simplicity we here restrict ourselves to the periodic setting with $I = [0, 2\pi]$ (opposite sides identified) and use Fourier series expansions. Note that analogous results can be obtained for non-periodic domains and other types of expansion systems with sufficient decay properties.

### 1.1. **Sobolev spaces for particles**

First, let us set up a basis for the one-particle space $\mathcal{H}^s(I^d) \subset \mathcal{L}^2(I^d)$. Here, we use the $d$-dimensional product of the one-dimensional system $\{\phi_k(x), k \in \mathbb{Z}\}$. An example are the trigonometric polynomials $\{\frac{1}{\sqrt{2\pi}} e^{ikx}, k \in \mathbb{Z},$

$x \in I\}$. We then define the $d$-dimensional multi-index $\mathbf{k} = (k_1, k_2, \ldots, k_d) \in \mathbb{Z}^d$, the coordinate vector $\mathbf{x} = (x_1, \ldots, x_d)$ and the associated $d$-dimensional basis functions

$$\phi_{\mathbf{k}}(\mathbf{x}) := \prod_{j=1}^{d} \phi_{k_j}(x_j). \tag{3}$$

We furthermore denote $|\mathbf{k}|_2 = (\sum_{i=1}^{d} k_i^2)^{1/2}$ and $|\mathbf{k}|_\infty = \max_{1\le j\le d} |k_j|$. Let us define (periodic) isotropic Sobolev spaces in $d$ dimensions *via* Fourier series, *i.e.* we classify functions *via* the decay of their Fourier coefficients. To this end, we set

$$\lambda(\mathbf{k}) := |\mathbf{k}|_2 \tag{4}$$

and define

$$\mathcal{H}^s(I^d) = \left\{ u(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} c_{\mathbf{k}} \phi_{\mathbf{k}}(\mathbf{x}) : \|u\|_{\mathcal{H}^s(I^d)} = \left( \sum_{\mathbf{k}\in\mathbb{Z}^d} (1+\lambda(\mathbf{k}))^{2s} \cdot |c_{\mathbf{k}}|^2 \right)^{1/2} \le c < \infty \right\}, \tag{5}$$

where $c_{\mathbf{k}} = \int_{I^d} \phi^*_{\mathbf{k}}(\mathbf{x}) u(\mathbf{x}) \mathrm{d}\mathbf{x}$ and $c$ is a constant which depends on $d$. Note that $\lambda(\mathbf{k})$ relates to the $\mathcal{H}^1$-seminorm, *i.e.*

$$-(\Delta u, u) = \sum_{\mathbf{k}\in\mathbb{Z}^d} c_{\mathbf{k}}^2 \lambda(\mathbf{k})^2.$$

Then, (5) is equivalent to the usual definition of the $\mathcal{H}^s$-norm, since

$$(1 + |\mathbf{k}|_2)^{2s} \simeq \sum_{|\alpha|\le s} |\mathbf{k}|_2^{2\alpha} \tag{6}$$

where the constants in the norm equivalence involve binomial coefficients with respect to $d$.

Based on the given one-particle basis (3) we now define a basis for many-particle spaces. We denote by $I^{d\cdot N} = I^d \times I^d \times \cdots \times I^d$ the $d \cdot N$ dimensional torus which is equivalent to the $d \cdot N$ dimensional cube where opposite faces are identified. We then have the $d \cdot N$-dimensional coordinates $\vec{\mathbf{x}} := (\mathbf{x}_1, \ldots, \mathbf{x}_N)$, where $\mathbf{x}_i \in I^d$. To this end, we first employ a tensor product construction and define the multi-indices $\vec{\mathbf{k}} = (\mathbf{k}_1, \ldots, \mathbf{k}_N) \in \mathbb{Z}^{d\cdot N}$ and the associated functions

$$\phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) := \prod_{i=1}^{N} \phi_{\mathbf{k}_i}(\mathbf{x}_i) = \left( \bigotimes_{i=1}^{N} \phi_{\mathbf{k}_i} \right)(\mathbf{x}_1, \ldots, \mathbf{x}_N). \tag{7}$$

They span the subspaces $V_{\vec{\mathbf{k}}} := \mathrm{span}(\phi_{\vec{\mathbf{k}}})$ whose union form[3] the space

$$V = \bigoplus_{\vec{\mathbf{k}}\in\mathbb{Z}^{d\cdot N}} V_{\vec{\mathbf{k}}}. \tag{8}$$

We then can uniquely represent any function $\psi$ from $V$ as

$$\psi(\vec{\mathbf{x}}) = \sum_{\vec{\mathbf{k}}\in\mathbb{Z}^{d\cdot N}} c_{\vec{\mathbf{k}}}\ \phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) \tag{9}$$

[3]Except for the completion with respect to a chosen Sobolev norm, $V$ is just the associated Sobolev space.

with coefficients $c_{\vec{\mathbf{k}}} \in \mathbb{C}$. For the specific choice of trigonometric polynomials $\{\frac{1}{\sqrt{2\pi}} e^{ikx}\}$ as one-dimensional basis on $I = [0, 2\pi]$ we get $c_{\vec{\mathbf{k}}} = (\frac{1}{2\pi})^{dN/2} \int_{I^{d\cdot N}} e^{-i\vec{\mathbf{k}}^T \vec{\mathbf{x}}} \psi(\vec{\mathbf{x}}) d\vec{\mathbf{x}}$. For a general $\mathcal{L}^2$-orthonormal basis $\{\phi_k\}$ we have of course $c_{\vec{\mathbf{k}}} = \int_{I^{d\cdot N}} \phi^*_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) \psi(\vec{\mathbf{x}}) d\vec{\mathbf{x}}$.

Now, starting from the one-particle space $\mathcal{H}^s(I^d)$ we build Sobolev spaces for many particles. Obviously there are many possibilities to generalize the concept of Sobolev spaces [2] from the one-particle case to higher dimensions. Two simple possibilities are the additive or multiplicative combination *i.e.* an arithmetic or geometric averaging of the frequencies for the different particles. We use the following definition that combines both possibilities. We denote

$$\lambda_{\text{mix}}(\vec{\mathbf{k}}) := \prod_{i=1}^{N} (1 + \lambda(\mathbf{k}_i)) \qquad \text{and} \qquad \lambda_{\text{iso}}(\vec{\mathbf{k}}) := 1 + \sum_{i=1}^{N} \lambda(\mathbf{k}_i). \tag{10}$$

Now, for $-\infty < t, l < \infty$, set

$$\mathcal{H}^{t,l}_{\text{mix}}((I^d)^N) = \left\{ u(\vec{\mathbf{x}}) = \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N}} c_{\vec{\mathbf{k}}} \phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) : \|u\|_{\mathcal{H}^{t,l}_{\text{mix}}((I^d)^N)} = \left( \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N}} \lambda_{\text{mix}}(\vec{\mathbf{k}})^{2t} \cdot \lambda_{\text{iso}}(\vec{\mathbf{k}})^{2l} \cdot |c_{\vec{\mathbf{k}}}|^2 \right)^{1/2} \leq c < \infty \right\} \tag{11}$$

with a constant $c$ which depends on $d$ and $N$.

Note that $\lambda_{\text{mix}}(\vec{\mathbf{k}})$ relates to the operator $\prod_{i=1}^{N}(1 - \Delta_i)$. It expresses the multiplicative combination of the $\mathcal{H}^1(I^d)$-norm of the one-particle space with a norm of the $N$-particle space which involves mixed derivatives. Furthermore, $\lambda_{\text{iso}}(\vec{\mathbf{k}})$ relates to the operator $I - \sum_{i=1}^{N} \Delta_i$ and creates directly an associated $\mathcal{H}^1((I^d)^N)$-norm for the $N$-particle space. A $t$- and $l$-times application of these operators leads together with the norm equivalence (6) to the corresponding multiplicative combination of the $\mathcal{H}^t(I^d)$-norm and the $\mathcal{H}^l((I^d)^N)$-norm, respectively.

The standard isotropic Sobolev spaces [2] as well as the Sobolev spaces of dominating mixed smoothness [69], both generalized to the $N$-particle case, are included here. They can be written as

$$\begin{aligned} \mathcal{H}^s((I^d)^N) &= \mathcal{H}^{0,s}_{\text{mix}}((I^d)^N) \\ &= \left\{ u(\vec{\mathbf{x}}) = \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N}} c_{\vec{\mathbf{k}}} \phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) : \|u\|_{\mathcal{H}^{s}_{\text{mix}}((I^d)^N)} = \left( \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N}} \lambda_{\text{iso}}(\vec{\mathbf{k}})^{2s} \cdot |c_{\vec{\mathbf{k}}}|^2 \right)^{1/2} \leq c < \infty \right\} \end{aligned} \tag{12}$$

and

$$\begin{aligned} \mathcal{H}^t_{\text{mix}}((I^d)^N) &= \mathcal{H}^{t,0}_{\text{mix}}((I^d)^N) \\ &= \left\{ u(\vec{\mathbf{x}}) = \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N}} c_{\vec{\mathbf{k}}} \phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) : \|u\|_{\mathcal{H}^{t,0}_{\text{mix}}((I^d)^N)} = \left( \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N}} \lambda_{\text{mix}}(\vec{\mathbf{k}})^{2t} \cdot |c_{\vec{\mathbf{k}}}|^2 \right)^{1/2} \leq c < \infty \right\}, \end{aligned}$$

respectively. Hence, the parameter $l$ from (11) governs the isotropic smoothness, whereas $t$ governs the mixed smoothness. Thus, the spaces $\mathcal{H}^{t,l}_{\text{mix}}$ give us a quite flexible framework for the study of problems in Sobolev spaces. Note that the relations $\mathcal{H}^t_{\text{mix}} \subset \mathcal{H}^t \subset \mathcal{H}^{t/N}_{\text{mix}}$ for $t \geq 0$ and $\mathcal{H}^{t/N}_{\text{mix}} \subset \mathcal{H}^t \subset \mathcal{H}^t_{\text{mix}}$ for $t \leq 0$ hold. See [69] and [48] for more information on the spaces $\mathcal{H}^t_{\text{mix}}$.

The spaces $\mathcal{H}^{t,l}_{\text{mix}}((I^d)^N)$ are a mixture of tensor-products [75] of one-particle Sobolev spaces: Let $t \in \mathbb{R}^+_0$, $l \in \mathbb{R}$, $t + l \geq 0$, $\mathbf{1} = (1, \ldots, 1)$ and $\mathbf{e_i} = (0, \ldots, 0, 1, 0, \ldots, 0)$ the $i$-th unit-vector in $\mathbb{R}^N$.

$$\mathcal{H}^{t,l}_{\text{mix}}((I^d)^N) := \mathcal{H}^{t\mathbf{1}+l\mathbf{e_1}}_{\text{mix}}((I^d)^N) \cap \cdots \cap \mathcal{H}^{t\mathbf{1}+l\mathbf{e_n}}_{\text{mix}}((I^d)^N), \tag{13}$$

where

$$\mathcal{H}^{\mathbf{k}}_{\text{mix}}((I^d)^N) := \mathcal{H}^{k_1}(T^d) \otimes \cdots \otimes \mathcal{H}^{k_n}(I^d).$$

This may easily be seen from the definition of the tensor-product *via* orthonormal systems and the intersection of spaces, compare [39, 42]. See also [49] for analogous constructions with more general boundary conditions. To prove $\mathcal{H}^s((I^d)^N) = \mathcal{H}^{0,s}_{\text{mix}}((I^d)^N)$ choose an orthogonal basis of $\mathcal{H}^s(I^d)$ and $\mathcal{L}^2(I^d)$, use periodic continuation to $\mathbb{R}^d$ and use the definition of the tensor-product *via* orthonormal systems [75]. Note that similar results hold for problems on $(I^d)^N$ with Dirichlet or Neumann boundary conditions and certain cases of mixed boundary conditions.

We have the simple norm equivalence

$$\sum_{i=1}^{N} \lambda(\mathbf{k}_i)^2 \simeq \max_{i=1,\ldots,N} \lambda(\mathbf{k}_i)^2$$

where the constant in the upper estimate involves a factor of $N$. This allows us to switch from $\lambda_{\text{iso}}(\vec{\mathbf{k}}) = 1 + \sum_{i=1}^{N} \lambda(\mathbf{k}_i)$ to

$$\lambda_{\text{iso}}(\vec{\mathbf{k}}) := 1 + \max_{i=1,\ldots,N} \lambda(\mathbf{k}_i). \tag{14}$$

With basically the same norm equivalence we can replace (4) by $\lambda(\mathbf{k}_i) = |\mathbf{k}_i|_\infty$. These changes in the definitions of $\lambda$ and $\lambda_{\text{iso}}$ result in the same spaces $\mathcal{H}^{t,l}_{\text{mix}}((I^d)^N)$ (with just a different $c$ in (11)). In the following we will work with these equivalent definitions since they simplify error estimates and complexity substantially.

## 1.2. **Regular sparse grids**

Now we are in the position to define finite-dimensional subspaces of $V$. First, we consider the conventional "full grid" space with respect to the discretization parameter $K \in \mathbb{N}$

$$V_{K,-\infty} := \bigoplus_{\substack{\vec{\mathbf{k}} \in \mathbb{Z}^{d \cdot N} \\ \lambda_{\text{iso}}(\vec{\mathbf{k}}) \le K+1}} V_{\vec{\mathbf{k}}} \tag{15}$$

with associated grid

$$\hat{\Omega}_{K,-\infty} := \{(\mathbf{k}_1, \ldots, \mathbf{k}_N) : \lambda_{\text{iso}}(\vec{\mathbf{k}}) \le K+1\}$$

in $\vec{\mathbf{k}}$-space. With the space $V_{K,-\infty}$ we obtain for $\psi \in \mathcal{H}^t((I^d)^N)$ with $s < t$ the standard error estimate

$$\inf_{V_{K,-\infty}} \|\psi - v\|_{\mathcal{H}^s} = \|\psi - \tilde{\psi}_{K,-\infty}\|_{\mathcal{H}^s} \le \|\psi - \psi_{K,-\infty}\|_{\mathcal{H}^s} = O((K+1)^{-(t-s)}) \cdot \|\psi\|_{\mathcal{H}^t} \tag{16}$$

where $\tilde{\psi}_{K,-\infty}$ denotes the best approximation in $V_{K,-\infty}$ with respect to the $\mathcal{H}^s$-norm and $\psi_{K,-\infty}$ denotes the interpolant of $\psi$ in $V_{K,-\infty}$, *i.e.* $\psi_{K,-\infty} = \sum_{\vec{\mathbf{k}} \in \hat{\Omega}_{K,-\infty}} c_{\vec{\mathbf{k}}} \phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})$. A straightforward calculation shows

$$|V_{K,-\infty}| = (2(K+1)+1)^{dN} = O((K+1)^{dN}). \tag{17}$$

Here we encounter the curse of dimensionality which renders such an approach impossible in practice for *e.g.* $d = 3, N > 1$.

Instead, we now define the sparse grid space with respect to the discretization parameter $K \in \mathbb{N}$ as

$$V_{K,0} := \bigoplus_{\substack{\vec{\mathbf{k}} \in \mathbb{Z}^{d \cdot N} \\ \lambda_{\text{mix}}(\vec{\mathbf{k}}) \le K+1}} V_{\vec{\mathbf{k}}} \tag{18}$$

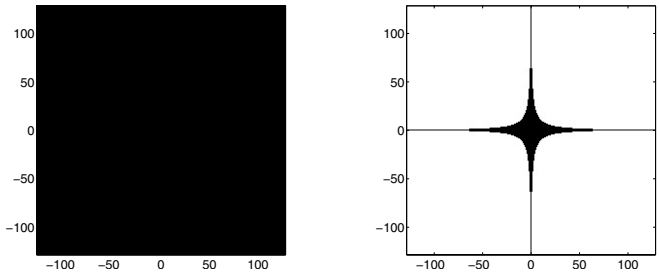


FIGURE 1. The (full) grid $\hat{\Omega}_{K,-\infty}$ (left) and the sparse grid $\hat{\Omega}_{K,0}$, for $d = 1$, $N = 2$ and $K = 128$.

with associated hyperbolic cross grid

$$\hat{\Omega}_{K,0} := \{(\mathbf{k}_1, \ldots, \mathbf{k}_N) : \lambda_{\text{mix}}(\vec{\mathbf{k}}) \leq K + 1\} \tag{19}$$

in $\vec{\mathbf{k}}$-space. This approach can be traced back at least to Korobov [5, 20], see also [72].

For the associated number of degrees of freedom we have the following lemma:

**Lemma 1.** The dimension of the sparse grid space/hyperbolic cross with respect to the discretization parameter $K \in \mathbb{N}$ is

$$|V_{K,0}| = O((K + 1)^d (\log(K + 1))^{N-1}). \tag{20}$$

A proof for the case $d = 1$, with $\lambda_{\text{mix}} = \prod_{i=1}^{N} \max(1, \lambda(\mathbf{k}_i))$, $\lambda(\mathbf{k}) = |\mathbf{k}|_\infty$ can be found in [78], see also [39, 54, 55, 79] and the related estimates in [9–12, 38]. A careful analysis is given in [21] for the case $d = 1$ where the estimate $2^N/(N - 1)! \cdot \det\Lambda$ for the constant in the $O$-notation is derived. Here, $\det\Lambda$ denotes the volume of the lattice unit cell $\Lambda$. It can be carried over to the case of general $d$ which results in an additional factor $d^{N-1}$ in the order constant.

In comparison to (17) we see that the number of degrees of freedom is now substantially reduced. The curse of dimension with respect to $N$ is only present in the $(\log K)^{N-1}$-term. Note however that the constant in the order estimate depends on $N$ and $d$. Figure 1 displays the index set for the full grid and the sparse grid for the case $d = 1, N = 2, K = 128$.

We now consider the error which is made by the approximation of a function $\psi \in V$ in the sparse grid subspace $V_{K,0}$. We have the following error estimate:

**Lemma 2.** Let $s < t$, $\psi \in \mathcal{H}^t_{\text{mix}}((I^d)^N)$. Let $\tilde{\psi}_{K,0}$ be the best approximation in $V_{K,0}$ with respect to the $\mathcal{H}^s$-norm and let $\psi_{K,0}$ be the interpolant of $\psi$ in $V_{K,0}$, *i.e.* $\psi_{K,0} = \sum_{\vec{\mathbf{k}} \in \hat{\Omega}_{K,0}} c_{\vec{\mathbf{k}}} \phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})$. Then, there holds

$$\inf_{V_{K,0}} \|\psi - v\|_{\mathcal{H}^s} = \|\psi - \tilde{\psi}_{K,0}\|_{\mathcal{H}^s} \leq \|\psi - \psi_{K,0}\|_{\mathcal{H}^s} = O((K + 1)^{-(t-s)}) \|\psi\|_{\mathcal{H}^t_{\text{mix}}}. \tag{21}$$

*Proof.* We have

$$\begin{aligned}
\|\psi - \psi_{K,0}\|^2_{\mathcal{H}^s} &= \| \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,0}} c_{\vec{\mathbf{k}}} \phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) \|^2_{\mathcal{H}^s} = \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,0}} |c_{\vec{\mathbf{k}}}|^2 \lambda_{\mathrm{iso}}(\vec{\mathbf{k}})^{2s} \qquad (22)\\
&= \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,0}} |c_{\vec{\mathbf{k}}}|^2 \lambda_{\mathrm{iso}}(\vec{\mathbf{k}})^{2s} \frac{\lambda_{\mathrm{mix}}(\vec{\mathbf{k}})^{2t}}{\lambda_{\mathrm{mix}}(\vec{\mathbf{k}})^{2t}}\\
&\leq \left( \max_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,0}} \frac{\lambda_{\mathrm{iso}}(\vec{\mathbf{k}})^{2s}}{\lambda_{\mathrm{mix}}(\vec{\mathbf{k}})^{2t}} \right) \left( \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,0}} |c_{\vec{\mathbf{k}}}|^2 \lambda_{\mathrm{mix}}(\vec{\mathbf{k}})^{2t} \right)\\
&\leq \left( \max_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,0}} \frac{\lambda_{\mathrm{iso}}(\vec{\mathbf{k}})^{2s}}{\lambda_{\mathrm{mix}}(\vec{\mathbf{k}})^{2t}} \right) \left( \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N}} |c_{\vec{\mathbf{k}}}|^2 \lambda_{\mathrm{mix}}(\vec{\mathbf{k}})^{2t} \right)\\
&= \max_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,0}} \frac{\lambda_{\mathrm{iso}}(\vec{\mathbf{k}})^{2s}}{\lambda_{\mathrm{mix}}(\vec{\mathbf{k}})^{2t}} \|\psi\|^2_{\mathcal{H}^t_{\mathrm{mix}}}.
\end{aligned}$$

Now, using the definition of the index set $\hat{\Omega}_{K,0}$ in (19) and evaluating the maximum in (23) we obtain the desired result (21). □

Hence there appears no loss in the order of approximation compared to the result (16) for the full grid approximation space provided that $\psi \in \mathcal{H}^t_{\mathrm{mix}}((I^d)^N)$. The involved degrees of freedom however are greatly reduced from $O(K^{dN})$ to $O(K^d \log(K)^{N-1})$.

Note that for the case of orthogonal basis systems other than $\{\frac{1}{\sqrt{2\pi}} \mathrm{e}^{ikx}\}$ we obtain analogous estimates. Then, the definition of $\lambda_{\mathrm{mix}}$ and $\lambda_{\mathrm{iso}}$ must be altered accordingly to express the respective decay factors for the corresponding derivatives. In case of a non-orthogonal system (with proper decay factors) additional Cauchy-Schwarz inequalities or related norm equivalences can be used to cope with the arising cross-terms, see *e.g.* [9, 10, 12, 38, 39].

### 1.3. **Optimized sparse grid spaces**

The use of the regular sparse grid space (18) substantially reduces the complexity in comparison to that of the full grid space if $\psi \in \mathcal{H}^t_{\mathrm{mix}}((I^d)^N)$. However there is still a curse of dimensionality present since the term $(\log K)^{N-1}$ grows exponentially with the number $N$ of particles. Even if $\log K$ is moderate in size, this limits the method to relatively small numbers $N$ of particles. The question is therefore if it is possible to get rid of this $(\log K)^{N-1}$-term. For conventional sparse grid methods (based on the one-dimensional hierarchical basis or wavelets with dyadic refinement) it was shown in [10, 12] that the sparse grid subspace selection can be justified from an optimization point of view: The task is to select a collection of subspaces $V_{\vec{\mathbf{k}}}$ for which the approximation error gets minimal for a given dimension of the approximation space (or *vice versa*). This global optimization problem can be cast into a binary knapsack problem which, after an embedding into a rational setting, can be easily solved. It turns out that for an optimal subspace collection just the subspaces $V_{\vec{\mathbf{k}}}$ have to be taken into account whose ratio of (an upper estimate of) the associated contribution to the error (benefit) *versus* the size of the subspace (cost) is larger than a prescribed threshold, see [10, 12] for further details. The benefit depends on the norm in which the approximation is sought. It turns out that the regular sparse grid is optimal in this respect for the $\mathcal{L}^p$-norms, $p \in [0, \infty]$. But if the error is measured in the $\mathcal{H}^1$-seminorm the optimization approach results in discrete subspaces which correspond to even more sparsified sparse grids. Their number of degrees of freedom is of the order $O(K)$ only, *i.e.* the above mentioned $(\log K)^{N-1}$-term is no longer present, provided that $\psi \in \mathcal{H}^{2,0}_{\mathrm{mix}}$. To this end we use dyadically refined, *i.e.* hierarchical piecewise linear hat functions as one-dimensional basis, for details see [10–12, 38]. The more general case of wavelet systems leads to analogous results which can be found in [39, 54].

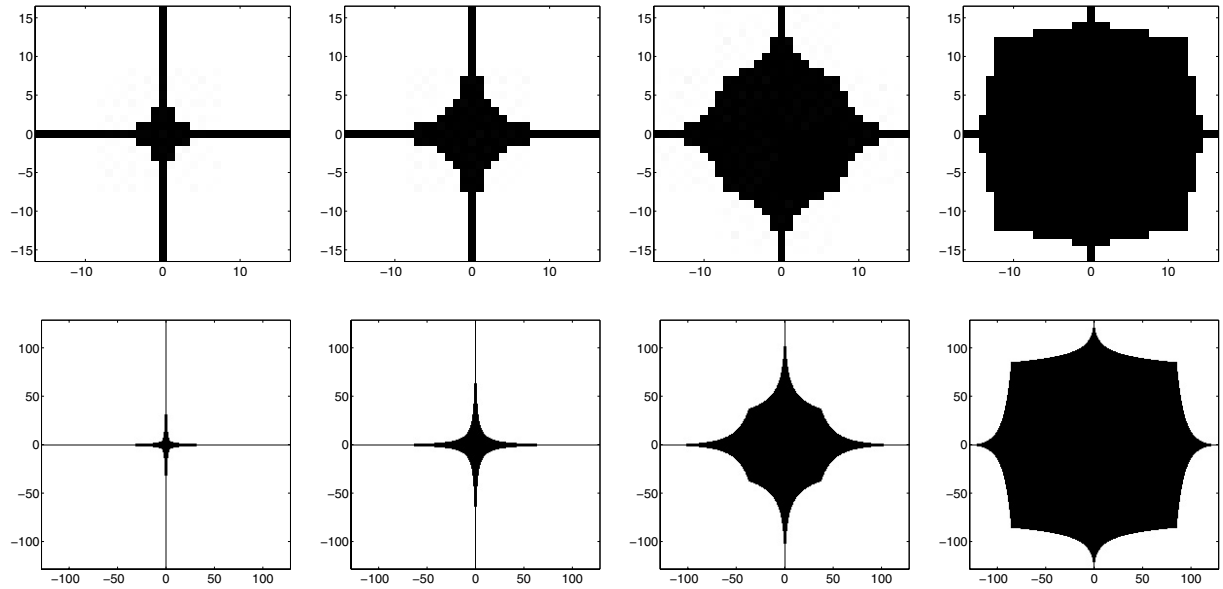

FIGURE 2. Index sets $\hat{\Omega}_{16,T}$ (above) and $\hat{\Omega}_{128,T}$ (below) for $T = 0.5, 0, -2, -10$ (from left to right), $d = 1, N = 2$; the hyperbolic cross corresponds to $T = 0$.

In the following we discuss an approach which is similar to such optimized sparse grid spaces and present the resulting complexities and approximation rates for the Fourier basis in the case of particle spaces. To this end, besides $K$, we introduce an additional parameter $T \in (-\infty, 1]$. We define the generalized sparse grid space

$$V_{K,T} := \bigoplus_{\substack{\vec{\mathbf{k}} \in \mathbb{Z}^{d \cdot N} \\ \lambda_{\mathrm{mix}}(\vec{\mathbf{k}}) \cdot \lambda_{\mathrm{iso}}(\vec{\mathbf{k}})^{-T} \leq K^{1-T}}} V_{\vec{\mathbf{k}}}$$

with associated generalized hyperbolic cross

$$\hat{\Omega}_{K,T} := \{(\mathbf{k}_1, \ldots, \mathbf{k}_N) : \lambda_{\mathrm{mix}}(\vec{\mathbf{k}}) \cdot \lambda_{\mathrm{iso}}(\vec{\mathbf{k}})^{-T} \leq K^{1-T}\}. \tag{23}$$

The parameter $T$ allows us to switch from the full grid case $T = -\infty$ to the sparse grid case $T = 0$ and also allows to create with $T \in (0, 1]$ subspaces of the hyperbolic cross space with further reduced complexities. Obviously, the inclusions $V_{K,T_1} \subset V_{K,T_2}$ for $T_1 \leq T_2$ hold. Figure 2 displays the index sets for various choices of $T$ for the case $d = 1, N = 2, K = 16$ and $K = 128$.

For the associated number of degrees of freedom we have the following result:

**Lemma 3.** The dimension of the generalized sparse grid space/hyperbolic cross with respect to the discretization parameter $K \in \mathbb{N}$ is

$$|V_{K,T}| = \begin{cases} O((K+1)^d) & \text{for } 0 < T \leq 1, \\ O((K+1)^d \cdot \log(K+1)^{N-1}) & \text{for } T = 0, \\ O((K+1)^{d \cdot \frac{T-1}{T/N-1}}) & \text{for } T < 0, \\ O((K+1)^{dN}) & \text{for } T = -\infty. \end{cases} \tag{24}$$

For a proof in the case $d = 1$ see the arguments in [54, 78] and compare [54, 55]. For our choice of $\lambda_{\mathrm{mix}}$ and $\lambda_{\mathrm{iso}}$ involving the maximum norm, these arguments can be straightforwardly carried over to the case of general $d$.

We see that the number of degrees of freedom is further reduced for the case $T \in (0,1]$. The curse of dimension with respect to $N$ which still was present in the $(\log(K+1))^{N-1}$-term in (20) has now completely disappeared. Note however that the constant in the order estimate still depends on $N$ and $d$.

The main question is now if for the advantageous case $T \in (0,1]$ the accuracy of the associated deteriorates or if it is maintained. The latter is indeed the case in certain situations. We have the following error estimate:

**Lemma 4.** Let $s < l+t$, $t \geq 0$, $\psi \in \mathcal{H}^{t,l}_{\text{mix}}((I^d)^N)$. Let $\tilde{\psi}_{K,T}$ be the best approximation in $V_{K,T}$ with respect to the $\mathcal{H}^s$-norm and let $\psi_{K,T}$ be the interpolant of $\psi$ in $V_{K,T}$, *i.e.* $\psi_{K,T} = \sum_{\vec{\mathbf{k}} \in \hat{\Omega}_{K,T}} c_{\vec{\mathbf{k}}} \phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})$. Then, there holds

$$\inf_{V_{K,T}} \|\psi - v\|_{\mathcal{H}^s} = \|\psi - \tilde{\psi}_{K,T}\|_{\mathcal{H}^s} \leq \|\psi - \psi_{K,T}\|_{\mathcal{H}^s} \leq \begin{cases} O((K+1)^{s-l-t+(Tt-s+l)\frac{N-1}{N-T}}) \cdot \|\psi\|_{\mathcal{H}^{t,l}_{\text{mix}}} & \text{for } T \geq \frac{s-l}{t}, \\ O((K+1)^{s-l-t}) \cdot \|\psi\|_{\mathcal{H}^{t,l}_{\text{mix}}} & \text{for } T \leq \frac{s-l}{t}. \end{cases} \tag{25}$$

*Proof.* We have, compare also [39, 54, 55],

$$\begin{aligned} \|\psi - \psi_{K,T}\|^2_{\mathcal{H}^s} &= \| \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,T}} c_{\vec{\mathbf{k}}} \phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) \|^2_{\mathcal{H}^s} = \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,T}} |c_{\vec{\mathbf{k}}}|^2 \lambda_{\text{iso}}(\vec{\mathbf{k}})^{2s} \qquad (26) \\ &= \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,T}} |c_{\vec{\mathbf{k}}}|^2 \lambda_{\text{iso}}(\vec{\mathbf{k}})^{2s} \frac{\lambda_{\text{iso}}(\vec{\mathbf{k}})^{2l}}{\lambda_{\text{iso}}(\vec{\mathbf{k}})^{2l}} \frac{\lambda_{\text{mix}}(\vec{\mathbf{k}})^{2t}}{\lambda_{\text{mix}}(\vec{\mathbf{k}})^{2t}} \\ &\leq \left( \max_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,T}} \frac{\lambda_{\text{iso}}(\vec{\mathbf{k}})^{2s}}{\lambda_{\text{iso}}(\vec{\mathbf{k}})^{2l}} \frac{1}{\lambda_{\text{mix}}(\vec{\mathbf{k}})^{2t}} \right) \left( \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,T}} |c_{\vec{\mathbf{k}}}|^2 \lambda_{\text{iso}}(\vec{\mathbf{k}})^{2l} \lambda_{\text{mix}}(\vec{\mathbf{k}})^{2t} \right) \\ &\leq \left( \max_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,T}} \frac{\lambda_{\text{iso}}(\vec{\mathbf{k}})^{2s}}{\lambda_{\text{iso}}(\vec{\mathbf{k}})^{2l}} \frac{1}{\lambda_{\text{mix}}(\vec{\mathbf{k}})^{2t}} \right) \left( \sum_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N}} |c_{\vec{\mathbf{k}}}|^2 \lambda_{\text{iso}}(\vec{\mathbf{k}})^{2l} \lambda_{\text{mix}}(\vec{\mathbf{k}})^{2t} \right) \\ &= \max_{\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N} \setminus \hat{\Omega}_{K,T}} \lambda_{\text{iso}}(\vec{\mathbf{k}})^{-2(l-s)} \lambda_{\text{mix}}(\vec{\mathbf{k}})^{-2t} \cdot \|\psi\|^2_{\mathcal{H}^{t,l}_{\text{mix}}} . \end{aligned}$$

Using the definition of the index set $\hat{\Omega}_{K,T}$ in (23) and evaluating the maximum in (27) we obtain the desired result (25). □

This type of estimate was already given for the case of a dyadically refined wavelet basis with $d = 1$ in [39, 54, 55]. It is a generalization of the energy-norm based sparse grid approach of [11, 12, 38] where the case $s = 1$, $t = 2$, $l = 0$ was considered using a hierarchical piecewise linear basis.

Altogether we have seen the following: for functions from $\mathcal{H}^{t,l}_{\text{mix}}((I^d)^N)$ the use of the generalized sparse grid space $V_{K,T}$ with $T \leq (s-l)/t$ leads to a significant reduction in the number of degrees of freedom compared to the full grid space $V_{K,-\infty}$ while the approximation order is preserved. Specifically, in the case $T \in (0,1]$ the dependence of the dimension of $V_{K,T}$ on $N$ with respect to the asymptotics in $K$ is completely removed while the order of approximation is for $T \leq (s-l)/t$ the same as in the full grid case.

Let us discuss some cases. For the standard Sobolev space $\mathcal{H}^{0,l}_{\text{mix}}$ (*i.e.* $t = 0$, $l = 2$) and the spaces $V_{K,T}$ with $T \geq -\infty$ the resulting approximation order is dependent on $T$ and dependent on the number of particles $N$. In particular the order even deteriorates with larger $T$. Note that for $T < 0$ the dimension of $V_{K,T}$ with respect to $K$ is exponentially dependent on $N$. This reflects the curse of dimensionality which makes problems in isotropic Sobolev spaces (12) intractable for higher values of $N$. For the standard Sobolev spaces of bounded

mixed derivatives $\mathcal{H}^{t,0}_{\text{mix}}$ (*i.e.* $t = 2,\ l = 0$) and the spaces $V_{K,T}$ with $T > \frac{s}{2}$ the resulting approximation order is dependent on $T$ and dependent on the number of particles $N$ whereas for $T \leq \frac{s}{2}$ the resulting order is independent of $T$ and $N$. Here, for $T \in (0, \frac{s}{2}]$ the dimension of $V_{K,T}$ according to (24) is independent of $N$. If we restrict the class of functions for example to $\mathcal{H}^{1,1}_{\text{mix}}$ (*i.e.* $t = 1,\ l = 1$) and measure the error in the $\mathcal{H}^1$-norm (*i.e.* $s = 1$) the approximation order is dependent on $N$ for all $T > 0$ and independent on $N$ and $T$ for all $T \leq 0$. In that case, for $T = 0$, the dependence of the dimension of $V_{K,T}$ on $N$ is only logarithmically. Note that in all cases the constants in the $O$-notation depend on $N$ and $d$.

We finally cast the estimates on the degrees of freedom and the associated error into a form which measures the error with respect to the involved degrees of freedom, *i.e.* the dimension $|V_{K,T}|$, and reach the following theorem:

**Theorem 5.** Let $\psi \in \mathcal{H}^{t,l}_{\text{mix}}((I^d)^N)$. Let $\tilde{\psi}_{K,T}$ be the best approximation in $V_{K,T}$ with respect to the $\mathcal{H}^s$-norm. Furthermore denote by $M$ the actual number of degrees of freedom of $V_{K,T}$, *i.e.* $M := |V_{K,T}|$. Consider the case $T \in (0, (s-l)/t]$. Then, there holds

$$\|\psi - \tilde{\psi}_{K,T}\|_{\mathcal{H}^s} \leq C(N,d) \cdot M^{-(l-s+t)/d} \cdot \|\psi\|_{\mathcal{H}^{t,l}_{\text{mix}}}$$

*Proof.* This is a simple consequence of the Lemmas 3 and 4. First, we put the definition $M = |V_{K,T}|$ into (24), *i.e.* into the relation $|V_{K,T}| \leq c_1(N,d) \cdot (K+1)^d$ and solve for $(K+1)^{-d}$. This results in $(K+1)^{-d} \leq c_1(N,d) M^{-1}$. We now plug this into (25), *i.e.* into the relation $\|\psi - \tilde{\psi}_{K,T}\|_{\mathcal{H}^s} \leq c_2(N,d) \cdot (K+1)^{-d((l-s)+t)/d} \cdot \|\psi\|_{\mathcal{H}^{t,l}_{\text{mix}}}$ and arrive at the desired result with $C(N,d) = c_1(N,d) \cdot c_2(N,d)$ where $c_1(N,d)$ denotes the order constant in (24) and $c_2(N,d)$ denotes the order constant in (25). □

Note finally that the constant $C$ still depends on $N$ and $d$. It is a difficult and tedious task to derive sharp estimates for the constants $c_1$ and $c_2$ and thus for $C$. So far we were able to show in one very special case, *i.e.* for $d = 1,\ s = 1,\ t = 2,\ l = 0$, with a piecewise linear basis on $I = [0,1]$ and vanishing boundary values of $\psi$, that $C(N,1)$ does not grow with $N$ but indeed decays exponentially. To be precise, we could show $C(N,1) \leq c \cdot N^2 \cdot 0.97515^N$, for further details see [38]. This behavior of the constant has of course to be compared with the behavior of $\|\psi\|_{\mathcal{H}^{2,0}_{\text{mix}}}$ which also depends on $N$. The derivation of precise estimates of the constants in the general cases and the study of the dependency of $\|\psi\|_{\mathcal{H}^{t,l}_{\text{mix}}}$ on $N$ is future work.

## 2. Antisymmetric sparse grids

So far we defined general sparse grid spaces and discussed their complexities and approximation properties. We now come back to the electronic Schrödinger equation (1). Note that in general an electronic wave function depends in addition to the positions $\mathbf{x}_i$ of the electrons also on their associated spin coordinates $\sigma_i \in \{-\frac{1}{2}, \frac{1}{2}\}$. Thus electronic wave functions are defined as

$$\Psi : (\mathbb{R}^d)^N \times \left\{-\frac{1}{2}, \frac{1}{2}\right\}^N \to \mathbb{R} : (\vec{\mathbf{x}}, \vec{\sigma}) \to \Psi(\vec{\mathbf{x}}, \vec{\sigma})$$

with spin coordinates $\vec{\sigma} = (\sigma_1, \ldots, \sigma_N)$. Furthermore, physically relevant eigenfunctions $\Psi$ obey the following two assumptions: first, elementary particles are indistinguishable from each other (fundamental principle of quantum mechanics). Second, no two electrons may occupy the same quantum state simultaneously (Pauli

exclusion principle)[4]. Thus, we consider only wave functions which are antisymmetric with respect to an arbitrary simultaneous permutation $P \in \mathcal{S}_N$ of the electron positions and spin variables, *i.e.* which fulfil

$$\Psi(P\vec{\mathbf{x}}, P\vec{\sigma}) = (-1)^{|P|}\Psi(P\vec{\mathbf{x}}, P\vec{\sigma}).$$

Here, $\mathcal{S}_N$ denotes the symmetric group. The permutation $P$ is a mapping $P : \{1, \ldots, N\} \to \{1, \ldots, N\}$ which translates to a permutation of the corresponding numbering of electrons and thus to a permutation of indices, *i.e.* we have $P(\mathbf{x}_1, \ldots, \mathbf{x}_N)^T := (\mathbf{x}_{P(1)}, \ldots, \mathbf{x}_{P(N)})^T$ and $P(\sigma_1, \ldots, \sigma_N)^T := (\sigma_{P(1)}, \ldots, \sigma_{P(N)})^T$. In particular, the symmetric group is of size $|\mathcal{S}_N| = N!$ and the expression $(-1)^{|P|}$ is equal to the determinant of the associated permutation matrix $\det P$.

Now, to a given spin vector $\vec{\sigma} \in \{-\frac{1}{2}, \frac{1}{2}\}^N$ we define the associated spatial component of the full wave function $\Psi$ by

$$\psi_{\vec{\sigma}} : (\mathbb{R}^d)^N \to \mathbb{R} : \vec{\mathbf{x}} \to \Psi(\vec{\mathbf{x}}, \vec{\sigma}).$$

Then, since there are $2^N$ possible different spin distributions $\vec{\sigma}$, the full Schrödinger equation, *i.e.* the eigenvalue problem $H\Psi = E\Psi$, decouples into $2^N$ eigenvalue problems for the $2^N$ associated spatial components $\psi_{\vec{\sigma}}$. Here, the spatial part $\psi_{\vec{\sigma}}$ to a given $\vec{\sigma}$ obeys the condition

$$\psi_{\vec{\sigma}}(P\vec{\mathbf{x}}) = (-1)^{|P|}\psi_{\vec{\sigma}}(P\vec{\mathbf{x}}), \quad \forall P \in \mathcal{S}_{\vec{\sigma}} := \{P \in \mathcal{S}_N \,:\, P\vec{\sigma} = \vec{\sigma}\}. \tag{27}$$

In particular, the minimal eigenvalue of all eigenvalue problems for the spatial components is equal to the minimal eigenvalue of the full eigenvalue problem. Moreover, the eigenfunctions of the full system can be composed by the eigenfunctions of the eigenvalue problems for the spatial parts.

Although there are $2^N$ possible different spin distributions $\vec{\sigma}$, the bilinear form $\langle\psi(P\cdot)|H|\psi(P\cdot)\rangle$ is invariant under all permutations $P \in \mathcal{S}_N$ of the position coordinates $\vec{\mathbf{x}}$. Thus it is sufficient to consider the eigenvalue problems which are associated to the spin vectors $\vec{\sigma}^{(N,S)} = (\sigma_1^{(N,S)}, \ldots, \sigma_N^{(N,S)})$ where the first $S$ electrons possess spin $-\frac{1}{2}$ and the remaining $N - S$ electrons possess spin $\frac{1}{2}$, *i.e.*

$$\sigma_j^{(N,S)} = \begin{cases} -\frac{1}{2} & \text{for } j \le S, \\ \frac{1}{2} & \text{for } j > S. \end{cases}$$

In particular, it is enough to solve only the $\lfloor N/2 \rfloor$ eigenvalue problems which correspond to the spin vectors $\vec{\sigma}^{(N,S)}$ with $S \le N/2$. For further details see [76]. Therefore, we consider in the following without loss of generality only spin distributions $\vec{\sigma}^{(N,S)} = (\sigma_1^{(N,S)}, \ldots, \sigma_N^{(N,S)})$. We set $\mathcal{S}_{(N,S)} := \mathcal{S}_{\vec{\sigma}^{(N,S)}}$. Note that there holds $|\mathcal{S}_{(N,S)}| = S!(N - S)!$.

Now we consider spaces of antisymmetric functions and their discrete sparse grid counterparts. The functions of the $N$-particle space $V$ from (8) which obey the anti-symmetry condition (27) for a given $\vec{\sigma}^{(N,S)}$ form a linear subspace $V^{\mathcal{A}^{(N,S)}}$ of $V$. We define the projection into this subspace, *i.e.* the antisymmetrization operator $\mathcal{A}^{(N,S)} : V \to V^{\mathcal{A}^{(N,S)}}$ by

$$\mathcal{A}^{(N,S)}\psi(\vec{\mathbf{x}}) := \frac{1}{S!(N-S)!} \sum_{P \in \mathcal{S}_{(N,S)}} (-1)^{|P|}\psi(P\vec{\mathbf{x}}) =: \psi^{(N,S)}(\vec{\mathbf{x}}). \tag{28}$$

[4]Fermions are elementary particles of half-integer spin. The Pauli principle states in general that no two identical fermions may occupy the same quantum state simultaneously. This does not apply to bosons which are elementary particles of integer spin.

For any basis function $\phi_{\vec{\mathbf{k}}}$ of our general $N$-particle space $V$ we then have

$$\begin{aligned}
\mathcal{A}^{(N,S)}\phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) &= \mathcal{A}^{(N,S)}\left(\left(\bigotimes_{i=1}^{S}\phi_{\mathbf{k}_i}\right)(\mathbf{x}_1,\ldots,\mathbf{x}_S)\left(\bigotimes_{i=S+1}^{N}\phi_{\mathbf{k}_i}\right)(\mathbf{x}_{S+1},\ldots,\mathbf{x}_N)\right)\\
&= \left(\mathcal{A}^{(S,S)}\bigotimes_{i=1}^{S}\phi_{\mathbf{k}_i}(\mathbf{x}_1,\ldots,\mathbf{x}_S)\right)\left(\mathcal{A}^{(N-S,N-S)}\bigotimes_{i=S+1}^{N}\phi_{\mathbf{k}_i}(\vec{\mathbf{x}}_{S+1},\ldots,\mathbf{x}_N)\right)\\
&= \left(\frac{1}{S!}\bigwedge_{i=1}^{S}\phi_{\mathbf{k}_i}(\mathbf{x}_1,\ldots,\mathbf{x}_S)\right)\left(\frac{1}{(N-S)!}\bigwedge_{i=S+1}^{N}\phi_{\mathbf{k}_i}(\mathbf{x}_{S+1},\ldots,\mathbf{x}_N)\right)\\
&= \frac{1}{S!}\begin{vmatrix}\phi_{\mathbf{k}_1}(\mathbf{x}_1) & \ldots & \phi_{\mathbf{k}_1}(\mathbf{x}_S)\\ \vdots & \ddots & \vdots\\ \phi_{\mathbf{k}_S}(\mathbf{x}_1) & \ldots & \phi_{\mathbf{k}_S}(\mathbf{x}_S)\end{vmatrix}\frac{1}{(N-S)!}\begin{vmatrix}\phi_{\mathbf{k}_{S+1}}(\mathbf{x}_{S+1}) & \ldots & \phi_{\mathbf{k}_{S+1}}(\mathbf{x}_N)\\ \vdots & \ddots & \vdots\\ \phi_{\mathbf{k}_N}(\mathbf{x}_{S+1}) & \ldots & \phi_{\mathbf{k}_N}(\mathbf{x}_N)\end{vmatrix}\\
&= \frac{1}{S!(N-S)!}\begin{vmatrix}\phi_{\mathbf{k}_1}(\mathbf{x}_1) & \ldots & \phi_{\mathbf{k}_1}(\mathbf{x}_S) & 0 & \ldots & 0\\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ \phi_{\mathbf{k}_S}(\mathbf{x}_1) & \ldots & \phi_{\mathbf{k}_S}(\mathbf{x}_S) & 0 & \ldots & 0\\ 0 & \ldots & 0 & \phi_{\mathbf{k}_{S+1}}(\mathbf{x}_{S+1}) & \ldots & \phi_{\mathbf{k}_{S+1}}(\mathbf{x}_N)\\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & \ldots & 0 & \phi_{\mathbf{k}_N}(\mathbf{x}_{S+1}) & \ldots & \phi_{\mathbf{k}_N}(\mathbf{x}_N)\end{vmatrix}\\
&= \frac{1}{S!(N-S)!}\sum_{P\in\mathcal{S}_{(N,S)}}(-1)^{|P|}\phi_{\vec{\mathbf{k}}}(P\vec{\mathbf{x}}) = \frac{1}{S!(N-S)!}\sum_{P\in\mathcal{S}_{(N,S)}}(-1)^{|P|}\phi_{P\vec{\mathbf{k}}}(\vec{\mathbf{x}}).
\end{aligned}$$

In other words, the classical product $\phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) := \prod_{i=1}^{N}\phi_{\mathbf{k}_i}(\mathbf{x}_i) = \left(\bigotimes_{i=1}^{N}\phi_{\mathbf{k}_i}\right)(\mathbf{x}_1,\ldots,\mathbf{x}_N)$ gets replaced by the product of two outer products

$$\frac{1}{S!}\bigwedge_{i=1}^{S}\phi_{\mathbf{k}_i}(\mathbf{x}_1,\ldots,\mathbf{x}_S) \quad \text{and} \quad \frac{1}{(N-S)!}\bigwedge_{i=S+1}^{N}\phi_{\mathbf{k}_i}(\mathbf{x}_{S+1},\ldots,\mathbf{x}_N)$$

that are associated to the two sets of coordinates and one-particle bases which are associated to the two spin values $-\frac{1}{2}$ and $\frac{1}{2}$. The outer product involves just the so-called Slater determinant [68], *i.e.*

$$\bigwedge_{i=1}^{N}\phi_{\mathbf{k}_i}(\mathbf{x}_1,\ldots,\mathbf{x}_N) = \begin{vmatrix}\phi_{\mathbf{k}_1}(\mathbf{x}_1) & \ldots & \phi_{\mathbf{k}_1}(\mathbf{x}_N)\\ \vdots & \ddots & \vdots\\ \phi_{\mathbf{k}_N}(\mathbf{x}_1) & \ldots & \phi_{\mathbf{k}_N}(\mathbf{x}_N)\end{vmatrix}.$$

The sequence $\left\{\mathcal{A}^{(N,S)}\phi_{\vec{\mathbf{k}}}\right\}_{\vec{\mathbf{k}}\in(\mathbb{Z}^d)^N}$ only forms a generating system of the antisymmetric subspace $V^{\mathcal{A}^{(N,S)}}$ and no basis since many functions $\mathcal{A}^{(N,S)}\phi_{\vec{\mathbf{k}}}$ are identical (up to the sign). This can be demonstrated using the simple example $N=2$ with $S=0$, *i.e.* with equal spin values $\frac{1}{2}$. In this case we have

$$\begin{aligned}
\mathcal{A}^{(2,0)}\phi_{(\mathbf{k}_1,\mathbf{k}_2)}(\mathbf{x}_1,\mathbf{x}_2) &= \frac{1}{2}\begin{vmatrix}\phi_{\mathbf{k}_1}(\mathbf{x}_1) & \phi_{\mathbf{k}_1}(\mathbf{x}_2)\\ \phi_{\mathbf{k}_2}(\mathbf{x}_1) & \phi_{\mathbf{k}_2}(\mathbf{x}_2)\end{vmatrix} = \frac{1}{2}(\phi_{\mathbf{k}_1}(\mathbf{x}_1)\phi_{\mathbf{k}_2}(\mathbf{x}_2) - \phi_{\mathbf{k}_1}(\mathbf{x}_2)\phi_{\mathbf{k}_2}(\mathbf{x}_1))\\
&= -(\frac{1}{2}\phi_{\mathbf{k}_2}(\mathbf{x}_1)\phi_{\mathbf{k}_1}(\mathbf{x}_2) - \phi_{\mathbf{k}_2}(\mathbf{x}_2)\phi_{\mathbf{k}_1}(\mathbf{x}_1)) = -\frac{1}{2}\begin{vmatrix}\phi_{\mathbf{k}_2}(\mathbf{x}_1) & \phi_{\mathbf{k}_2}(\mathbf{x}_2)\\ \phi_{\mathbf{k}_1}(\mathbf{x}_1) & \phi_{\mathbf{k}_1}(\mathbf{x}_2)\end{vmatrix}\\
&= -\mathcal{A}^{(2,0)}\phi_{(\mathbf{k}_2,\mathbf{k}_1)}(\mathbf{x}_1,\mathbf{x}_2)
\end{aligned}$$

*i.e.* two different basis functions of $V$ are mapped by the antisymmetrization operator to the same function with just a different sign.

We can gain a basis for the antisymmetric subspace $V^{\mathcal{A}^{(N,S)}}$ if we restrict the sequence $\left\{\mathcal{A}^{(N,S)}\phi_{\vec{\mathbf{k}}}\right\}_{\vec{\mathbf{k}}\in(\mathbb{Z}^d)^N}$ properly. This can be done in many different ways. A possible orthonormal basis $\mathcal{B}^{(N,S)}$ for $V^{\mathcal{A}^{(N,S)}}$ is given with help of

$$\Phi^{(N,S)}_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) := \left(\frac{1}{\sqrt{S!}}\bigwedge_{i=1}^{S}\phi_{\mathbf{k}_i}\left(\mathbf{x}_1,\ldots,\mathbf{x}_S\right)\right)\left(\frac{1}{\sqrt{(N-S)!}}\bigwedge_{i=S+1}^{N}\phi_{\mathbf{k}_i}\left(\mathbf{x}_{S+1},\ldots,\mathbf{x}_N\right)\right) \tag{29}$$

as follows:

$$\mathcal{B}^{(N,S)} := \left\{\Phi^{(N,S)}_{\vec{\mathbf{k}}} : \vec{\mathbf{k}}\in(\mathbb{Z}^d)^N,\ \mathbf{k}_1<\ldots<\mathbf{k}_S\ \wedge\ \mathbf{k}_{S+1}<\ldots<\mathbf{k}_N\right\} \tag{30}$$

where the relation $<$ is defined as

$$\mathbf{k}_i<\mathbf{k}_j\ :\Leftrightarrow\ \exists\alpha\in\{1,\ldots,d\}:\ \mathbf{k}_{i,(\alpha)}<\mathbf{k}_{j,(\alpha)}\ \wedge\ \forall\beta\in\{1,\ldots,\alpha-1\}:\ \mathbf{k}_{i,(\beta)}\le\mathbf{k}_{j,(\beta)}. \tag{31}$$

We then can define the antisymmetric subspace $V^{\mathcal{A}^{(N,S)}}$ of $V$ as

$$V^{\mathcal{A}^{(N,S)}} = \bigoplus_{\substack{\vec{\mathbf{k}}\in\mathbb{Z}^{d\cdot N}\\ \mathbf{k}_1<\ldots<\mathbf{k}_S\ \wedge\ \mathbf{k}_{S+1}<\ldots<\mathbf{k}_N}} V_{\vec{\mathbf{k}}}, \tag{32}$$

where we denote from now on $V_{\vec{\mathbf{k}}} := \operatorname{span}(\Phi^{(N,S)}_{\vec{\mathbf{k}}})$. Any function $\psi^{\mathcal{A}^{(N,S)}}$ from $V^{\mathcal{A}^{(N,S)}}$ can now uniquely be represented as

$$\psi^{\mathcal{A}^{(N,S)}}(\vec{\mathbf{x}}) = \sum_{\substack{\vec{\mathbf{k}}\in\mathbb{Z}^{d\cdot N}\\ \mathbf{k}_1<\ldots<\mathbf{k}_S\ \wedge\ \mathbf{k}_{S+1}<\ldots<\mathbf{k}_N}} c_{\vec{\mathbf{k}}}\ \Phi^{(N,S)}_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) \tag{33}$$

with coefficients $c_{\vec{\mathbf{k}}} = \int_{I^{d\cdot N}}\Phi^{(N,S)^*}_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})\psi^{\mathcal{A}^{(N,S)}}(\vec{\mathbf{x}})\mathrm{d}\vec{\mathbf{x}}$.

Now we are in the position to consider finite-dimensional subspaces of $V^{\mathcal{A}^{(N,S)}}$. To this end, with the parameters $K\in\mathbb{N}$ and $T\in(-\infty,1]$, we define the the generalized antisymmetric sparse grid space

$$V^{\mathcal{A}^{(N,S)}}_{K,T} := \bigoplus_{\substack{\vec{\mathbf{k}}\in\mathbb{Z}^{d\cdot N}\\ \lambda_{\mathrm{mix}}(\vec{\mathbf{k}})\lambda_{\mathrm{iso}}(\vec{\mathbf{k}})^{-T}\le(K+1)^{1-T}\\ \mathbf{k}_1<\ldots<\mathbf{k}_S\ \wedge\ \mathbf{k}_{S+1}<\ldots<\mathbf{k}_N}} V_{\vec{\mathbf{k}}}$$

with associated antisymmetric generalized hyperbolic cross

$$\hat{\Omega}^{\mathcal{A}^{(N,S)}}_{K,T} := \{(\mathbf{k}_1,\ldots,\mathbf{k}_N):\lambda_{\mathrm{mix}}(\vec{\mathbf{k}})\lambda_{\mathrm{iso}}(\vec{\mathbf{k}})^{-T}\le(K+1)^{1-T},\mathbf{k}_1<\ldots<\mathbf{k}_S\ \wedge\ \mathbf{k}_{S+1}<\ldots<\mathbf{k}_N\}.$$

We then can uniquely represent any function $\psi^{\mathcal{A}^{(N,S)}}_{K,T}$ from $V^{\mathcal{A}^{(N,S)}}_{K,T}$ in a finite series

$$\psi^{\mathcal{A}^{(N,S)}}_{K,T}(\vec{\mathbf{x}}) = \sum_{\substack{\vec{\mathbf{k}}\in\mathbb{Z}^{d\cdot N}\\ \lambda_{\mathrm{mix}}(\vec{\mathbf{k}})\lambda_{\mathrm{iso}}(\vec{\mathbf{k}})^{-T}\le(K+1)^{1-T}\\ \mathbf{k}_1<\ldots<\mathbf{k}_S\ \wedge\ \mathbf{k}_{S+1}<\ldots<\mathbf{k}_N}} c_{\vec{\mathbf{k}}}\ \Phi^{(N,S)}_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})$$

with coefficients $c_{\vec{\mathbf{k}}} = \int_{I^{d\cdot N}}\Phi^{(N,S)^*}_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})\psi^{\mathcal{A}^{(N,S)}}_{K,T}(\vec{\mathbf{x}})\mathrm{d}\vec{\mathbf{x}}$. Again, the parameter $T$ allows us to switch from the antisymmetric full grid case $T=-\infty$ to the antisymmetric sparse grid case $T=0$ and allows to create with $T\in(0,1]$ subspaces of the antisymmetric hyperbolic cross space with further reduced complexities. Obviously,

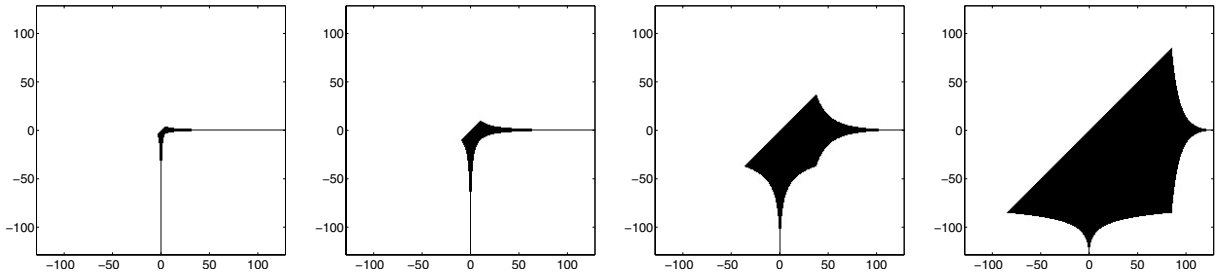

FIGURE 3. The antisymmetric sparse grids $\hat{\Omega}_{K,T}^{\mathcal{A}^{(2,1)}}$ (above) and $\hat{\Omega}_{K,T}^{\mathcal{A}^{(2,0)}}$ (below) for $T = 0.5$, $0, -2, -10$ where $d = 1$, $N = 2$ and $K = 128$.

the inclusions $V_{K,T_1}^{\mathcal{A}^{(N,S)}} \subset V_{K,T_2}^{\mathcal{A}^{(N,S)}}$ for $T_1 \leq T_2$ hold. Figure 3 displays the index sets for various choices of $T$ for the case $d = 1, N = 2, K = 128$. Note that in this special case with $N = 2$ the index set $\hat{\Omega}_{K,T}^{\mathcal{A}^{(2,1)}}$ is just identical to $\hat{\Omega}_{K,T}$ from Figure 2.

For the associated number of degrees of freedom we have the following lemma:

**Lemma 6.** The dimension of the generalized antisymmetric sparse grid space/hyperbolic cross with respect to the discretization parameter $K \in \mathbb{N}$ is

$$|V_{K,T}^{\mathcal{A}^{(N,S)}}| \leq \frac{1}{S! \cdot (N-S)!} \cdot |V_{K,T}|. \tag{34}$$

*Proof.* This is a straightforward consequence of the definition of the antisymmetrization operator (28) and the definition (30) and (31). □

We see that the order for the dimension of the different spaces with respect to $K$ stays for all different cases of $T$ the same as in the classical case of Lemma 3 without antisymmetry. However the constant is now reduced by the factor $1/(S!(N-S)!)$. This is a substantial improvement which allows to treat larger numbers of particles in the antisymmetric case.

Also with respect to the achieved accuracies the order in $K$ does not change when we switch to the antisymmetric case. Additionally the involved order constants do not change. We have in the antisymmetric case the same error estimates as in the as in the classical case of Lemma 4.

**Lemma 7.** Let $s < l + t$, $t \geq 0$, $\psi^{\mathcal{A}^{(N,S)}} = \mathcal{A}^{(N,S)}\psi \in \mathcal{A}^{(N,S)}\mathcal{H}_{\text{mix}}^{t,l}((I^d)^N)$. Let $\tilde{\psi}_{K,T}^{\mathcal{A}^{(N,S)}}$ be the best approximation in $V_{K,T}^{\mathcal{A}^{(N,S)}}$ with respect to the $\mathcal{H}^s$-norm and let $\psi_{K,T}^{\mathcal{A}^{(N,S)}}$ the interpolant of $\psi^{\mathcal{A}^{(N,S)}}$ in $V_{K,T}^{\mathcal{A}^{(N,S)}}$, *i.e.* $\psi_{K,T}^{\mathcal{A}^{(N,S)}} = \sum_{\vec{\mathbf{k}} \in \hat{\Omega}_{K,T}^{\mathcal{A}^{(N,S)}}} c_{\vec{\mathbf{k}}} \Phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})$. Then, there holds

$$\begin{aligned} \inf_{V_{K,T}^{\mathcal{A}^{(N,S)}}} \|\psi^{\mathcal{A}^{(N,S)}} - v\|_{\mathcal{H}^s} &\leq \|\psi^{\mathcal{A}^{(N,S)}} - \psi_{K,T}^{\mathcal{A}^{(N,S)}}\|_{\mathcal{H}^s} \\ &\leq \begin{cases} O((K+1)^{s-l-t+(Tt-s+l)\frac{N-1}{N-T}}) \cdot \|\psi^{\mathcal{A}^{(N,S)}}\|_{\mathcal{H}_{\text{mix}}^{t,l}} & \text{for } T \geq \frac{s-l}{t}, \\ O((K+1)^{s-l-t}) \cdot \|\psi^{\mathcal{A}^{(N,S)}}\|_{\mathcal{H}_{\text{mix}}^{t,l}} & \text{for } T \leq \frac{s-l}{t}. \end{cases} \end{aligned} \tag{35}$$

*Proof.* We derive along the lines of the proof of Lemma 4

$$\|\psi^{\mathcal{A}^{(N,S)}} - \psi_{K,T}^{\mathcal{A}^{(N,S)}}\|_{\mathcal{H}^s}^2 \leq \max_{\substack{\vec{\mathbf{k}} \in \mathbb{Z}^{d \cdot N} \\ \mathbf{k}_1 < \ldots < \mathbf{k}_S \wedge \mathbf{k}_{S+1} < \ldots < \mathbf{k}_N}} \setminus \hat{\Omega}_{K,T}^{\mathcal{A}^{(N,S)}} \lambda_{\text{iso}}(\vec{\mathbf{k}})^{-2(l-s)} \lambda_{\text{mix}}(\vec{\mathbf{k}})^{-2t} \cdot \|\psi\|_{\mathcal{H}_{\text{mix}}^{t,l}}^2 .$$

Using the definition of the index set $\hat{\Omega}_{K,T}^{\mathcal{A}^{(N,S)}}$ and evaluating the maximum we obtain (35). Note that the maximization gives the same order for the estimates as in Lemma 4. However in certain cases a slightly smaller maximum is achieved due to the constraint in (30). □

We again cast the estimates on the degrees of freedom and the associated error into a form which measures the error with respect to the involved degrees of freedom, *i.e.* the dimension $|V_{K,T}^{\mathcal{A}^{(N,S)}}|$, and reach the following theorem:

**Theorem 8.** Let $\psi^{\mathcal{A}^{(N,S)}} \in \mathcal{A}^{(N,S)}\mathcal{H}_{\text{mix}}^{t,l}((I^d)^N)$. Let $\tilde{\psi}_{K,T}^{\mathcal{A}^{(N,S)}}$ be the best approximation in $V_{K,T}^{\mathcal{A}^{(N,S)}}$ with respect to the $\mathcal{H}^s$-norm. Furthermore denote by $M_{\mathcal{A}^{(N,S)}}$ the actual number of degrees of freedom of $V_{K,T}^{\mathcal{A}^{(N,S)}}$, *i.e.* $M_{\mathcal{A}^{(N,S)}} := |V_{K,T}^{\mathcal{A}^{(N,S)}}|$. Consider the case $T \in (0, (s-l)/t]$. Then, there holds

$$\|\psi^{\mathcal{A}^{(N,S)}} - \tilde{\psi}_{K,T}^{\mathcal{A}^{(N,S)}}\|_{\mathcal{H}^s} \leq C(N,d)/(S! \cdot (N-S)!) \cdot M_{\mathcal{A}^{(N,S)}}^{-(l-s+t)/d} \cdot \|\psi^{\mathcal{A}^{(N,S)}}\|_{\mathcal{H}_{\text{mix}}^{t,l}}$$

*Proof.* Same as for Theorem 5 but with the results of Lemmas 6 and 7. □

Note finally that the constant still depends on $N$ and $d$. However in contrast to the non-antisymmetric case we now obtained an improvement by the factor $1/(S!(N-S)!)$. It is nevertheless still a tedious task to derive sharp estimates for the involved constants. The derivation of precise estimates of the constants and the study of the dependency of $\|\psi^{\mathcal{A}^{(N,S)}}\|_{\mathcal{H}_{\text{mix}}^{t,l}}$ on $N$ remains to be done.

## 3. Regularity of the solution of Schrödinger's equation

So far we introduced various optimized sparse grid spaces for the discretization of particle problems and discussed their dimensions and approximation properties. We furthermore carried these discretization techniques over to the case of antisymmetric wave functions. Here, the optimal complexity orders with respect to the number $N$ of particles for both, the dimension of the resulting discrete spaces and their associated approximation rates depended on the degree $s$ of the Sobolev-norm in which we measure the approximation error and the degrees $t$ and $l$ of isotropic and anisotropic smoothness, respectively, which was assumed to hold for the continuous wave function.

We now return to the electronic Schrödinger problem (1) and invoke our general theory for this special case. To this end, let us recall a major result from [77]. There, Yserentant showed that an antisymmetric solution of the electronic Schrödinger equation with $d = 3$ possesses basically $\mathcal{H}_{\text{mix}}^{1,1}$- or $\mathcal{H}_{\text{mix}}^{1/2,1}$-regularity. To be precise, he showed that an eigenfunction $\psi_{\vec{\sigma}}$ to a given spin distribution $\vec{\sigma}^{(N,S)}$ has certain square integrable mixed derivatives of order up to $S+1$

$$\int \lambda_{\text{mix},-\frac{1}{2}}^2(\vec{\mathbf{k}})\lambda_{\text{iso}}^2(\vec{\mathbf{k}})|\hat{\psi}(\vec{\mathbf{k}})|^2 \mathrm{d}\vec{\mathbf{k}} < \infty$$

with respect to the coordinates $\vec{\mathbf{x}}_1, \ldots, \vec{\mathbf{x}}_S$ and certain square integrable mixed derivatives of order up to $N-S+1$

$$\int \lambda_{\text{mix},\frac{1}{2}}^2(\vec{\mathbf{k}})\lambda_{\text{iso}}^2(\vec{\mathbf{k}})|\hat{\psi}(\vec{\mathbf{k}})|^2 \mathrm{d}\vec{\mathbf{k}} < \infty$$

with respect to the coordinates $\vec{\mathbf{x}}_{S+1}, \ldots, \vec{\mathbf{x}}_N$. Here $\hat{\psi}$ denotes the Fourier transform of $\psi$ and $\lambda_{\text{mix},\sigma}$ is given by

$$\lambda_{\text{mix},\sigma}(\vec{\mathbf{k}}) := \begin{cases} \prod_{i=1}^{S}(1+\lambda(\mathbf{k}_i)) & \text{for } \sigma = -\frac{1}{2}, \\ \prod_{i=S+1}^{N}(1+\lambda(\mathbf{k}_i)) & \text{for } \sigma = \frac{1}{2}. \end{cases}$$

Therefore, a full antisymmetric solution, *i.e.* $S = 0$ or $S = N$, possesses $\mathcal{H}^{1,1}_{\text{mix}}$-regularity. Furthermore, in the case of an arbitrary chosen $1 \leq S \leq N$, the inequality[5]

$$\int \lambda_{\text{mix}}(\vec{\mathbf{k}})\lambda^2_{\text{iso}}(\vec{\mathbf{k}})|\hat{\psi}(\vec{\mathbf{k}})|^2 \mathrm{d}\vec{\mathbf{k}} \leq \frac{1}{2} \sum_{\sigma \in \{-\frac{1}{2},\frac{1}{2}\}} \int \lambda^2_{\text{mix},\sigma}(\vec{\mathbf{k}})\lambda^2_{\text{iso}}(\vec{\mathbf{k}})|\hat{\psi}(\vec{\mathbf{k}})|^2 \mathrm{d}\vec{\mathbf{k}}$$

holds due to the elementary relation

$$\prod_{i=1}^{N} |1 + \lambda(\mathbf{k}_i)| \leq \frac{1}{2} \prod_{i=1}^{S} |1 + \lambda(\mathbf{k}_i)|^2 + \frac{1}{2} \prod_{i=S+1}^{N} |1 + \lambda(\mathbf{k}_i)|^2$$

and thus any partial antisymmetric wavefunction possesses at least $\mathcal{H}^{\frac{1}{2},1}_{\text{mix}}$-regularity.

The main argument to derive this fact is a Hardy type inequality, see [77] for details. Note that the derivation of (1) was done for the whole space $I = \mathbb{R}$. However it is easy to see that the arguments of [77] can be carried over also to the case of a finite domain.

Let us now consider the case of a full antisymmetric solution, *i.e.* the case $S = 0$ or $S = N$, and the resulting approximation rate and complexity in more detail. If we measure the approximation error in the $\mathcal{H}^1$-norm, we obtain from Lemma 7 with $s = 1$ and $t = l = 1$ ($\mathcal{H}^{1,1}_{\text{mix}}$-regularity) the approximation order $O((K+1)^{-1+T\cdot\frac{N-1}{N-T}})$ for $T \geq 0$ and $O((K+1)^{-1}) = O(K^{-1})$ for $T \leq 0$. In particular, for the choice $T = 0$ we have a rate of $O(K)$. Here, however, the dimension of the associated antisymmetric sparse grid space $V^{\mathcal{A}^{(N,S)}}_{K,T}$ is of the order $O((K+1)^d \log(K+1)^{N-1})$, see Lemmas 3 and 6, *i.e.* there is still an exponential dependency of the costs on $N$ with respect to $K$ in the $\log(K+1)^{N-1}$-term. If we now choose $T$ such that $0 < T < 1$ the cost complexity drops to $O((K+1)^d)$. The approximation rate then is of the order $O(K^{-1+T\cdot\frac{N-1}{N-T}}) = O(K^{-1+T-T\cdot\frac{1-T}{N-T}})$ which results in $O(K^{-1+T})$ for $N \to \infty$. Thus for example for the choice $T = 0.05$ we would obtain a rate of the order $O(K^{-0.95})$ with $O(K^d)$ involved degrees of freedom[6]. In an analog way we can argue for the partial antisymmetric case where we have for an arbitrary chosen $0 \leq S \leq N$ at least $\mathcal{H}^{\frac{1}{2},1}_{\text{mix}}$-regularity of the associated wave function. Then, with $0 < T < 1/2$ we obtain for $N \to \infty$ an approximation rate of the order $O(K^{-1/2+T/2})$ whereas the associated cost complexity is again $O((K+1)^d)$. For our example of $T = 0.05$ we would obtain at least a rate of the order $O(K^{-0.475})$ with $O(K^d)$ involved degrees of freedom.

## 4. Discretization, assembly of the eigenvalue system and solution

We now consider the assembly of the discrete system matrix which is associated to the generalized antisymmetric sparse grid space $V^{\mathcal{A}^{(N,S)}}_{K,T}$ from (32). As basis functions we use the corresponding antisymmetric $\Phi^{(N,S)}_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})$ from (29) with one-particle basis functions $\phi_{\mathbf{k}}(\mathbf{x})$ of the type (3) in a Galerkin discretization of (1) for whose indices $\vec{\mathbf{k}}$ the condition

$$\vec{\mathbf{k}} \in \mathbb{Z}^{d\cdot N}, \quad \lambda_{\text{mix}}(\vec{\mathbf{k}})\lambda_{\text{iso}}(\vec{\mathbf{k}})^{-T} \leq (K+1)^{1-T}, \quad \mathbf{k}_1 < \ldots < \mathbf{k}_S \wedge \mathbf{k}_{S+1} < \ldots < \mathbf{k}_N$$

holds. To this end, we fix $N > 0$ and $1 \leq S \leq N$ and omit for reasons of simplicity the indices $S$ and $N$ in the following. We also omit the indices $K$ and $T$ when they are clear from the context.

[5] Yserentant uses here $\lambda_{\text{iso}}(\vec{\mathbf{k}}) := 1 + \sum_{i=1}^{N} \lambda(\mathbf{k}_i)^2$ and $\lambda(\mathbf{k}_i) := |\mathbf{k}_i|_2$ which is (up to constants) equivalent to our definitions (10) and (14), respectively.

[6] This of course only holds asymptotically for $N$ large enough. It is not surprising since, for fixed $N$, it holds $K \log K^{N-1} \leq c \cdot K^{1-\varepsilon}$ for any $\varepsilon > 0$ with sufficiently large $K$. Also note that the constant in the order estimate still depends on $N$ and $d$.

To each pair of indices $\vec{\mathbf{k}}, \vec{\mathbf{l}}$ and associated functions $\Phi_{\vec{\mathbf{k}}}, \Phi_{\vec{\mathbf{l}}}$ we obtain one entry in the stiffness matrix, *i.e.*

$$A_{\vec{\mathbf{k}},\vec{\mathbf{l}}} := \langle \Phi_{\vec{\mathbf{k}}} | H | \Phi_{\vec{\mathbf{l}}} \rangle = \int \Phi^*_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) H \Phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) \, \mathrm{d}\vec{\mathbf{x}}. \tag{36}$$

Next we assume that we use orthogonal one-particle basis functions $\phi_{\mathbf{k}}(\mathbf{x})$. Then, we have also orthogonality of the antisymmetric many-particle basis functions, *i.e.* $\langle \Phi_{\vec{\mathbf{k}}} | \Phi_{\vec{\mathbf{k}}} \rangle = \int \Phi^*_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) \Phi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) \, \mathrm{d}\vec{\mathbf{x}} = \delta_{\vec{\mathbf{k}},\vec{\mathbf{l}}}$. We then can take advantage of the well-known Slater-Condon rules [17, 68, 71]. Consequently, quite a few entries of the system matrix are zero and the remaining non-zero entries can be put together from the values of certain $d$- and $2d$-dimensional integrals. This will be explained in more detail in the following.

Let us introduce an operator $O_1 := \sum_{i=1}^{N} o_1(i)$ acting on the $d \cdot N$-dimensional wavefunction where the one-particle operator $o_1(i)$, albeit applied to a $d \cdot N$-dimensional wavefunction, only acts on its $i$-th ($d$-dimensional) component. In particular, for problem (1) we set here

$$o_1(i) = -\frac{1}{2}\Delta_{\mathbf{x}_i} - \begin{cases} -\sum_{j=1}^{N_{nuc}} Z_j |\mathbf{x}_i - \mathbf{R}_j|_2, & d = 1, \\ -\sum_{j=1}^{N_{nuc}} Z_j \log(|\mathbf{x}_i - \mathbf{R}_j|_2), & d = 2, \\ \sum_{j=1}^{N_{nuc}} \frac{Z_j}{|\mathbf{x}_i - \mathbf{R}_j|_2}, & d = 3, \end{cases}$$

where $\mathbf{R}_j \in \mathbb{R}^d$ denotes the position of the $j$-th nucleus. We furthermore introduce the operator

$$o_1 = -\frac{1}{2}\Delta_{\mathbf{x}} - \begin{cases} -\sum_{j=1}^{N_{nuc}} Z_j |\mathbf{x} - \mathbf{R}_j|_2, & d = 1, \\ -\sum_{j=1}^{N_{nuc}} Z_j \log(|\mathbf{x} - \mathbf{R}_j|_2), & d = 2, \\ \sum_{j=1}^{N_{nuc}} \frac{Z_j}{|\mathbf{x} - \mathbf{R}_j|_2}, & d = 3, \end{cases} \qquad (= o_1(1))$$

which applies directly to a $d$-dimensional function.

Let us in addition introduce an operator $O_2 := \sum_{i<j=1}^{N} o_2(i,j)$ acting on the $d \cdot N$-dimensional wavefunction where the two-particle operator $o_2(i,j)$, albeit applied to a $d \cdot N$-dimensional wavefunction, only acts on its two $i$-th and $j$-th $d$-dimensional components. In particular, for problem (1) we set here

$$o_2(i,j) = \begin{cases} -|\mathbf{x}_i - \mathbf{x}_j|_2, & d = 1, \\ -\log|\mathbf{x}_i - \mathbf{x}_j|_2, & d = 2, \\ 1/|\mathbf{x}_i - \mathbf{x}_j|_2, & d = 3. \end{cases}$$

We furthermore introduce the operator

$$o_2 = \begin{cases} -|\mathbf{x} - \mathbf{y}|_2, & d = 1, \\ -\log|\mathbf{x} - \mathbf{y}|_2, & d = 2, \\ 1/|\mathbf{x} - \mathbf{y}|_2, & d = 3, \end{cases}$$

which applies directly to a $2d$-dimensional function. We then see that $o_2(i,j) = o_2(j,i)$. Now we can write the Schrödinger operator $H$ as

$$H = O_1 + O_2. \tag{37}$$

To shorten notation let us finally introduce

$$h(\mathbf{k}_1, \mathbf{k}_2) := \int \phi^*_{\mathbf{k}_1}(\mathbf{x}) o_1 \phi_{\mathbf{k}_2}(\mathbf{x})\, \mathrm{d}\mathbf{x}, \tag{38}$$

$$G(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) := \int \phi^*_{\mathbf{k}_1}(\mathbf{x}) \phi^*_{\mathbf{k}_2}(\mathbf{y}) o_2 \phi_{\mathbf{k}_3}(\mathbf{x}) \phi_{\mathbf{k}_4}(\mathbf{y})\, \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y}, \tag{39}$$

$$g(\mathbf{k}_1, \mathbf{k}_2) := G(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_1, \mathbf{k}_2), \tag{40}$$

$$\tilde{g}(\mathbf{k}_1, \mathbf{k}_2) := G(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_2, \mathbf{k}_1). \tag{41}$$

In particular, there holds $G(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4) = G(\mathbf{k}_2, \mathbf{k}_1, \mathbf{k}_4, \mathbf{k}_3)$.

We are now ready to state the so-called Slater-Condon rules (adapted to our setting). Due to the orthogonality of the one-dimensional particle basis used and the fact that the Schrödinger operator $H$ is made up of sums of one- and two-particle operators as explained above, we just have to distinguish four different cases for the indices $\vec{\mathbf{k}} = (\mathbf{k}_1, \ldots, \mathbf{k}_N)$ and $\vec{\mathbf{l}} = (\mathbf{l}_1, \ldots, \mathbf{l}_N)$ of a pair of test and trial function in (36). We consider the case where all one-particle basis functions involved in $\Phi_{\vec{\mathbf{k}}}$ and $\Phi_{\vec{\mathbf{l}}}$ are pairwise the same, the case where just one one-particle basis function differs, the case where two one-particle basis functions differ and the case where more than two one-particle basis functions differ[7]. A longer tedious calculation then shows the following:

**Case 1.** Identical one-particle basis functions, *i.e.* $\Phi_{\vec{\mathbf{k}}}, \Phi_{\vec{\mathbf{l}}}$ where $\vec{\mathbf{k}} = (\mathbf{k}_1, \ldots, \mathbf{k}_N) = \vec{\mathbf{l}} = (\mathbf{l}_1, \ldots, \mathbf{l}_N)$:

$$\begin{aligned}\langle \Phi_{\vec{\mathbf{k}}} | H | \Phi_{\vec{\mathbf{k}}} \rangle = \; & \sum_{j=1}^{N} h(\mathbf{k}_j, \mathbf{k}_j) + \frac{1}{2} \sum_{n \neq m}^{S} \left( g(\mathbf{k}_n, \mathbf{k}_m) - \tilde{g}(\mathbf{k}_n, \mathbf{k}_m) \right) \\ & + \frac{1}{2} \sum_{S < n \neq m}^{N} \left( g(\mathbf{k}_n, \mathbf{k}_m) - \tilde{g}(\mathbf{k}_n, \mathbf{k}_m) \right) + \sum_{n=1}^{S} \sum_{m=S+1}^{N} g(\mathbf{k}_n, \mathbf{k}_m).\end{aligned}$$

**Case 2.** Just one pair of one-particle basis functions is different, *i.e.* $\Phi_{\vec{\mathbf{k}}}, \Phi_{\vec{\mathbf{l}}}$ where $!\exists\, \mu \in \{1, \ldots, N\}$, $\mathbf{k}_\mu \neq \mathbf{l}_\mu$ with $\vec{\mathbf{l}} = (\mathbf{k}_1, \ldots, \mathbf{k}_{\mu-1}, \mathbf{l}_\mu, \mathbf{k}_{\mu+1}, \ldots, \mathbf{k}_N)$. We then have to distinguish two different subcases:

$$\mu \leq S: \qquad \langle \Phi_{\vec{\mathbf{k}}} | H | \Phi_{\vec{\mathbf{l}}} \rangle = h(\mathbf{k}_\mu, \mathbf{l}_\mu) + \sum_{n \neq \mu}^{S} \Big( G(\mathbf{k}_n, \mathbf{k}_\mu, \mathbf{k}_n, \mathbf{l}_\mu) - G(\mathbf{k}_n, \mathbf{k}_\mu, \mathbf{l}_\mu, \mathbf{k}_n) \Big) + \sum_{n=S+1}^{N} G(\mathbf{k}_\mu, \mathbf{k}_n, \mathbf{l}_\mu, \mathbf{k}_n),$$

$$\mu > S: \qquad \langle \Phi_{\vec{\mathbf{k}}} | H | \Phi_{\vec{\mathbf{l}}} \rangle = h(\mathbf{k}_\mu, \mathbf{l}_\mu) + \sum_{S < n \neq \mu}^{N} \Big( G(\mathbf{k}_n, \mathbf{k}_\mu, \mathbf{k}_n, \mathbf{l}_\mu) - G(\mathbf{k}_n, \mathbf{k}_\mu, \mathbf{l}_\mu, \mathbf{k}_n) \Big) + \sum_{n=1}^{S} G(\mathbf{k}_\mu, \mathbf{k}_n, \mathbf{l}_\mu, \mathbf{k}_n).$$

**Case 3.** Two pairs of one-particle basis functions are different, *i.e.* $\Phi_{\vec{\mathbf{k}}}, \Phi_{\vec{\mathbf{l}}}$ where $!\exists\, \mu_1 < \mu_2 \in \{1, \ldots, N\}$, $\mathbf{k}_{\mu_1} \neq \mathbf{l}_{\mu_1} \wedge \mathbf{k}_{\mu_2} \neq \mathbf{l}_{\mu_2}$ with $\vec{\mathbf{l}} = (\mathbf{k}_1, \ldots, \mathbf{k}_{\mu_1 - 1}, \mathbf{l}_{\mu_1}, \mathbf{k}_{\mu_1+1}, \ldots, \mathbf{k}_{\mu_2-1}, \mathbf{l}_{\mu_2}, \mathbf{k}_{\mu_2+1}, \ldots, \mathbf{k}_N)$. Again, we have to distinguish two different subcases:

$$\begin{aligned}\mu_1 < \mu_2 \leq S,\ S < \mu_1 < \mu_2: \qquad & \langle \Phi_{\vec{\mathbf{k}}} | H | \Phi_{\vec{\mathbf{l}}} \rangle = G(\mathbf{k}_{\mu_1}, \mathbf{k}_{\mu_2}, \mathbf{k}_{\mu_1}, \mathbf{l}_{\mu_2}) - G(\mathbf{k}_{\mu_1}, \mathbf{k}_{\mu_2}, \mathbf{l}_{\mu_2}, \mathbf{l}_{\mu_1}), \\ \mu_1 \leq S < \mu_2: \qquad & \langle \Phi_{\vec{\mathbf{k}}} | H | \Phi_{\vec{\mathbf{l}}} \rangle = G(\mathbf{k}_{\mu_1}, \mathbf{k}_{\mu_2}, \mathbf{k}_{\mu_1}, \mathbf{l}_{\mu_2}).\end{aligned}$$

[7]Before the Slater-Condon rules can be used the two determinants $\Phi^{(N,S)}_{\vec{\mathbf{k}}}$ and $\Phi^{(N,S)}_{\vec{\mathbf{l}}}$ must be arranged in maximum coincidence.

**Case 4.** More than two pairs of one-particle basis functions are different, *i.e.* $\Phi_{\vec{\mathbf{k}}}, \Phi_{\vec{\mathbf{l}}} \in \mathcal{B}^{(N,S)}$ where $\exists \mu_1 < \mu_2 < \mu_3 \in \{1, \ldots, N\}$, $\mathbf{k}_{\mu_1} \neq \mathbf{l}_{\mu_1} \wedge \mathbf{k}_{\mu_2} \neq \mathbf{l}_{\mu_2} \wedge \mathbf{k}_{\mu_3} \neq \mathbf{l}_{\mu_3}$:

$$\langle \Phi_{\vec{\mathbf{k}}} | H | \Phi_{\vec{\mathbf{l}}} \rangle = 0.$$

We thus see that for all the index pairs $\vec{\mathbf{k}}, \vec{\mathbf{l}}$ which are of Case 4 we directly obtain zero entries in the system matrix. Furthermore, from the other three cases we see that the non-zero entries can be put together from just the values of the $d$-dimensional integrals $h, g, \tilde{g}$ and the $(2 \cdot d)$-dimensional integrals $G$. Here it is advisable to compute and store these data on the fly (in *e.g.* a hash table) when needed for the first time and to reuse it when needed again in the computation of another matrix entry.

Note that we assumed so far only orthogonality of the one-particle basis functions but made no specific choice for them yet. This depends also on how we deal with the domain of the Schrödinger equation in the discretization process. To this end, recall that the solution of (1) lives on the whole $\mathbb{R}^{d \cdot N}$. To obtain a practically manageable situation we have to impose further restrictions to Schrödinger's equation, its domain and boundary conditions. Here mainly two possibilities exist. First, we could decide for a periodic setting. Then we may use the finite domain $([-a_1/2, a_1/2) \times \cdots \times [-a_d/2, a_d/2))^N$ and we may employ the one-particle basis functions $\phi_{\mathbf{k}}(\mathbf{x}) := \prod_{j=1}^{d} \phi_{k_j}(x_j)$ from (3) with

$$\phi_{k_j}(x_j) = \frac{1}{\sqrt{a_j}} \mathrm{e}^{2\pi i k_j x_j / a_j} \tag{42}$$

with $\mathbf{k} \in \mathbb{Z}^d$ which fulfil periodic boundary conditions. Within this setting the one-particle Coulomb operator and the two-particle Coulomb operator for $d = 3$ become $\sum_{\mathbf{L}} \sum_{j=1}^{N_{nuc}} \frac{-Z_j}{|\mathbf{x}-\mathbf{R_j}-\mathbf{L}|_2}$ and $\sum_{\mathbf{L}} \frac{1}{|\mathbf{x}-\mathbf{y}-\mathbf{L}|_2}$, respectively. Here $\mathbf{L}$ are lattice vectors that map the one-particle unit cell $[-a_1/2, a_1/2) \times \cdots \times [-a_d/2, a_d/2)$ into its periodic images. The problem however is that this lattice sum does not converge and has to be replaced by the well-known Ewald potential, see [23] for a further discussion.

An alternative is to simply choose a sufficiently large finite domain, *e.g.* $([-a_1/2, a_1/2) \times \cdots \times [-a_d/2, a_d/2))^N$, such that the solution for a tiny given molecular system which is put into the middle of the domain nearly vanishes at the boundary[8]. Thus it is reasonable to also truncate the Coulomb interaction potentials at a properly chosen distance $D$. This way the above-mentioned periodicity problem and the Ewald summation is avoided. Note that this approach of course introduces an error which depends on the size $\mathbf{a} = (a_1, \ldots, a_d)^T$ of the domain and the truncation parameter $D$. This error can made arbitrary small by enlarging $\mathbf{a}$ and $D$, respectively, (which however also enlarges the amount of basis functions needed) and has to be balanced properly with the discretization error.

In our numerical experiments we will follow this approach and will use for reasons of simplicity here the one-particle basis functions (42) with $a = a_1 = \cdots = a_d$ to build up the antisymmetric $N$-particle basis functions (29) and the associated antisymmetric generalized sparse grid spaces $V_{K,T}^{\mathcal{A}^{(N,S)}}$ from it. Then, the one- and two-particle integrals $h, g, \tilde{g}, G$ from (38–41) which are necessary to set up the entries of the system matrix $A$ can be computed analytically for $d = 1$ and $d = 3$.

First we consider the $d$-dimensional integral expressions related to the kinetic energy operator $-\frac{1}{2}\Delta_{\mathbf{x}}$. They can be analytically computed for $d = 1$, $d = 2$ and $d = 3$ as

$$\langle \phi_{\mathbf{k}} | -\frac{1}{2}\Delta | \phi_{\mathbf{l}} \rangle = \frac{1}{2} \int_{I^d} \nabla \phi_{\mathbf{k}}(\mathbf{x}) \cdot \nabla \phi_{\mathbf{l}}(\mathbf{x}) \mathrm{d}\mathbf{x} = \frac{2\pi^2}{a^2} |\mathbf{k}|_2^2 \delta_{\mathbf{k},\mathbf{l}},$$

where $I = [-a/2, a/2)$.

[8]Note that there is also the possibility to use a smaller finite domain and to employ absorbing boundary conditions [3,25,53,74]. However they are more involved than just simple homogeneous Dirichlet conditions.

Let us now consider the terms related to the truncated Coulomb potential

$$v_D(\mathbf{r}) = \begin{cases} v(\mathbf{r}), & |\mathbf{r}|_2 \le D, \\ 0, & \text{otherwise}, \end{cases}$$

where

$$v(\mathbf{r}) = \begin{cases} -|\mathbf{r}|_2, & d = 1, \\ -\log|\mathbf{r}|_2, & d = 2, \\ 1/|\mathbf{r}|_2, & d = 3. \end{cases}$$

Here, we have to compute the $d$-dimensional integrals

$$\int_{I^d} \phi^*_{\mathbf{k}}(\mathbf{x}) v_D(\mathbf{x} - \mathbf{R}) \phi_{\mathbf{l}}(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{43}$$

with $\mathbf{R} \in I^d$ and the $2d$-dimensional integrals

$$\int_{I^d} \int_{I^d} \phi^*_{\mathbf{k}}(\mathbf{x}) \phi^*_{\mathbf{m}}(\mathbf{y}) v_D(\mathbf{x} - \mathbf{y}) \phi_{\mathbf{l}}(\mathbf{x}) \phi_{\mathbf{n}}(\mathbf{y}) \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y}. \tag{44}$$

If we assume that the wavefunction is zero outside the set $\{\mathbf{x} : |\mathbf{x}|_2 < \tilde{D}\}$ then we choose a truncation radius $D = 2\tilde{D}$ and the size parameter $a = 2D$.

In the case $d = 3$ the 6-dimensional integral (44) can be written with the help of a coordinate transformation $\mathbf{r} = \mathbf{x} - \mathbf{y}$ and $\tilde{\mathbf{r}} = \frac{1}{2}(\mathbf{x} + \mathbf{y})$ in the form of a 3-dimensional integral

$$\begin{aligned} \int_{I^3} \int_{I^3} \phi^*_{\mathbf{k}}(\mathbf{x}) \phi^*_{\mathbf{m}}(\mathbf{y}) v_D(\mathbf{x} - \mathbf{y}) \phi_{\mathbf{l}}(\mathbf{x}) \phi_{\mathbf{n}}(\mathbf{y}) \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y} &= \frac{1}{a^6} \int_{I^3} \int_{I^3} \mathrm{e}^{-i\frac{2\pi}{a}(\mathbf{k}-\mathbf{l})^T \mathbf{x}} \mathrm{e}^{-i\frac{2\pi}{a}(\mathbf{m}-\mathbf{n})^T \mathbf{y}} v_D(\mathbf{x} - \mathbf{y}) \mathrm{d}\mathbf{x}\, \mathrm{d}\mathbf{y} \\ &= \frac{1}{a^3} \int_{(2I)^3} \mathrm{e}^{-i\frac{\pi}{a}(\mathbf{k}-\mathbf{l}+\mathbf{n}-\mathbf{m})^T \mathbf{r}} v_D(\mathbf{r}) \mathrm{d}\mathbf{r} \delta_{\mathbf{k}-\mathbf{l}+\mathbf{m},\mathbf{n}} \end{aligned} \tag{45}$$

and for $\mathbf{n} = \mathbf{k} - \mathbf{l} + \mathbf{m}$ we obtain

$$\begin{aligned} \frac{1}{a^3} \int_{(2I)^3} \mathrm{e}^{-i\frac{2\pi}{a}(\mathbf{k}-\mathbf{l})^T \mathbf{r}} v_D(\mathbf{r}) \mathrm{d}\mathbf{r} &= \frac{1}{a^3} \int_{|\mathbf{r}|_2 \le D} \mathrm{e}^{-i\frac{2\pi}{a}(\mathbf{k}-\mathbf{l})^T \mathbf{r}} v_D(\mathbf{r}) \mathrm{d}\mathbf{r} \\ &= \begin{cases} \frac{4\pi}{a^3} \frac{1}{k^2} (1 - \cos(kD)), & k > 0 \\ \frac{4\pi}{a^3} \frac{D}{2}, & k = 0 \end{cases} \end{aligned}$$

with the help of spherical coordinates, where $k = \frac{2\pi}{a}|\mathbf{k} - \mathbf{l}|_2$. Analogously the 3-dimensional integral (43) can be computed for $\mathbf{R} = 0$ by

$$\begin{aligned} \frac{1}{a^3} \int_{I^3} \mathrm{e}^{-i\frac{2\pi}{a}(\mathbf{k}-\mathbf{l})^T \mathbf{r}} v_D(\mathbf{r}) \mathrm{d}\mathbf{r} &= \frac{1}{a^3} \int_{|\mathbf{r}|_2 \le D} \mathrm{e}^{-i\frac{2\pi}{a}(\mathbf{k}-\mathbf{l})^T \mathbf{r}} v_D(\mathbf{r}) \mathrm{d}\mathbf{r} \\ &= \begin{cases} \frac{4\pi}{a^3} \frac{1}{k^2} (1 - \cos(kD)), & k > 0, \\ \frac{4\pi}{a^3} \frac{D}{2}, & k = 0. \end{cases} \end{aligned}$$

In the case $d = 2$ the 4-dimensional integral (44) for $\mathbf{n} = \mathbf{k} - \mathbf{l} + \mathbf{m}$ and the 2-dimensional integral (43) for $\mathbf{R} = 0$ can be reduced with the help of polar coordinates to one-dimensional integrals

$$\frac{1}{a^2} \int_{|\mathbf{r}|_2 < D} \mathrm{e}^{-i\frac{2\pi}{a}(\mathbf{k}-\mathbf{l})^T \mathbf{r}} v_D(\mathbf{r}) \mathrm{d}\mathbf{r} = -\frac{2\pi}{a^2} \int_0^D r \ln(r) J_0(kr) \mathrm{d}r,$$

where $J_0$ denotes the zero-order Bessel function of first kind and $k = \frac{2\pi}{a}|\mathbf{k}-\mathbf{l}|_2$. This integral is related to the Hankel transformation or so-called Fourier-Bessel transformation and we compute it numerically.

In the case $d = 1$ we compute the integral (44) analogously to the case $d = 3$ for $\mathbf{n} = \mathbf{k} - \mathbf{l} + \mathbf{m}$ and the integral (43) for $\mathbf{R} = 0$ by

$$\frac{1}{a}\int_{|\mathbf{r}|_2<D} \mathrm{e}^{-i\frac{2\pi}{a}(\mathbf{k}-\mathbf{l})^T\mathbf{r}} v_D(\mathbf{r})\mathrm{d}\mathbf{r} = \begin{cases} -\frac{2}{a}\frac{1}{k^2}(kD\sin(kD)+\cos(kD)-1), & k>0, \\ -\frac{D^2}{a}, & k=0, \end{cases}$$

where $k = \frac{2\pi}{a}|\mathbf{k}-\mathbf{l}|_2$. With the help of these formulae it is then straightforward to compute the entries of the system matrix $A$.

For the solution of the resulting discrete eigenvalue problem we invoke a parallelized conventional Lanczos method taken from the software package SLEPc [47] which is based on parallel software package PETSc [6]. Note that here also other solution approaches are possible with improved complexities, like multigrid-type methods [13,14,56,62] which however still need to be carried over to the setting of our generalized antisymmetric sparse grids.

## 5. Numerical experiments

We now turn to the results of numerical experiments with our new discretization method using generalized antisymmetric sparse grids. In the following, we choose the finite domain $([-a_1/2, a_1] \times \cdots \times [-a_d/2, a_d/2])^N$ with fixed $(a_1, \ldots, a_d)$, $a_i = a$, and restrict ourselves to the Schrödinger operator (1) where the involved Coulomb potential is truncated at a distance $D$. As one-particle basis functions we employ $\phi_{\mathbf{k}}(\mathbf{x}) := \prod_{j=1}^d \phi_{k_j}(x_j)$ from (3) with $\phi_{k_j}(x_j)$ from (42). Note that an estimate like (35) for the accuracy of an eigenfunction relates to an analogous estimate for the eigenvalue by means of the relation $|E - E^{app}| \le 4 \cdot \|\psi - \psi^{app}\|^2_{\mathcal{L}^2}$ where $E$ and $\psi$ denote the exact minimal eigenvalue and associated eigenfunction of $H$, respectively, and $E^{app}$ and $\psi^{app}$ denote finite-dimensional Galerkin approximations in arbitrary subspaces, see also [76]. Then, with Lemma 7, we would obtain with $s = 0, l = 1, t = 1$ and $S = 0$ the estimate

$$|E^{\mathcal{A}^{(N,0)}} - E^{\mathcal{A}^{(N,0)}}_{K,T}| \le 4 \cdot \|\psi^{\mathcal{A}^{(N,0)}} - \psi^{\mathcal{A}^{(N,0)}}_{K,T}\|^2_{\mathcal{L}^2} \le O((K+1)^{2\cdot(-2+(T+1)\frac{N-1}{N-T})}) \cdot \|\psi^{\mathcal{A}^{(N,0)}}\|^2_{\mathcal{H}^{1,1}_{\mathrm{mix}}}$$

for the case $d = 3$ and we see that the eigenvalues are in general much better approximated than the eigenfunctions. For example, for $T = 0$, this would result in a (squared) rate of the order $-4 + 2(N-1)/N$ which is about $-4$ for small numbers of $N$ but gets $-2$ for $N \to \infty$. Analogous arguments can be made for the cases $T = 0.25$ as well as for the case $S = N/2$ with reduced regularity involving the values $t = 1/2, l = 1$.

First we consider the case of one-dimensional particles. For varying numbers $N$ of particles we study the behavior of the discrete energy $E$, *i.e.* the smallest eigenvalue of the associated system matrix $A$, as $K$ increases. Here, we use the generalized antisymmetric sparse grid space $V^{\mathcal{A}^{(N,S)}}_{K,T}$ from (32) and focus on the two cases $T = 0$ and $T = 0.25$ with either $S = 0$ or $S = \lfloor N/2 \rfloor$. Tables 1–4 give the obtained results. Here, $\#A$ denotes the number of the non-zero matrix entries. Note that we have, besides the zero entries due to case four of the Slater-Condon rules, additional zero entries due to (45). The memory requirements of our algorithm are of the order $O(\#A)$. Furthermore, $\Delta E$ denotes the difference of the obtained values of $E$ and $\varepsilon$ denotes the quotient of the values of $\Delta E$ for two successive rows in the table. Thus, $\varepsilon$ indicates the convergence rate of the discretization error.

We see that the method is indeed convergent if the value for $K$ increases. Also the convergence rate gets improved for rising $K$. Moreover, for smaller values of $N$ but large values of $K$ we observe very fast convergence. We conjecture that this might also be the case for larger numbers $N$ of electrons with sufficient resolutions $K$ larger than the ones reached in our tables. Note that we have for our problem no $\mathcal{H}^{1,1}_{\mathrm{mix}}$- or $\mathcal{H}^{1/2,1}_{\mathrm{mix}}$-regularity statement at hand like in the three-dimensional case and we therefore do not know what convergence rate may be expected. However it seems that there is more regularity present which might cause the increase in the

TABLE 1. $d = 1$, $S = 0$, $a = 20$, $D = 10$, $T = 0$, $Z = N$.

| $N$ | $K$ | $M_{\mathcal{A}^{(N,S)}}$ | $\#A$ | $E$ | $\Delta E$ | $\epsilon$ |
|---|---|---|---|---|---|---|
| 1 | 2 | 5 | 17 | 1.416 581 | | |
| 1 | 4 | 9 | 49 | 0.951 454 | 4.651 265 e-01 | |
| 1 | 8 | 17 | 161 | 0.815 773 | 1.356 814 e-01 | 3.428 078 |
| 1 | 16 | 33 | 577 | 0.808 738 | 7.035 095 e-03 | 19.286 364 |
| 1 | 32 | 65 | 2177 | 0.808 620 | 1.183 037 e-04 | 59.466 418 |
| 1 | 64 | 129 | 8449 | 0.808 617 | 3.005 205 e-06 | 39.366 257 |
| 1 | 128 | 257 | 33 281 | 0.808 617 | 9.119 886 e-08 | 32.952 216 |
| 1 | 256 | 513 | 132 097 | 0.808 617 | 2.859 494 e-09 | 31.893 357 |
| 2 | 2 | 4 | 12 | 9.477 604 | | |
| 2 | 4 | 9 | 45 | 6.822 791 | 2.654 813 e+00 | |
| 2 | 8 | 26 | 232 | 4.263 177 | 2.559 614 e+00 | 1.037 193 |
| 2 | 16 | 67 | 1015 | 3.249 651 | 1.013 526 e+00 | 2.525 455 |
| 2 | 32 | 176 | 4544 | 2.820 979 | 4.286 719 e-01 | 2.364 339 |
| 2 | 64 | 431 | 18 815 | 2.760 278 | 6.070 105 e-02 | 7.062 018 |
| 2 | 128 | 1030 | 76 924 | 2.758 640 | 1.638 270 e-03 | 37.051 932 |
| 2 | 256 | 2407 | 309 655 | 2.758 532 | 1.080 520 e-04 | 15.161 865 |
| 2 | 512 | 5529 | 1 240 989 | 2.758 526 | 6.265 451 e-06 | 17.245 684 |
| 4 | 16 | 4 | 12 | 36.622 042 | | |
| 4 | 32 | 32 | 256 | 24.566 288 | 1.205 575 e+01 | |
| 4 | 64 | 158 | 2508 | 17.794 734 | 6.771 554 e+00 | 1.780 353 |
| 4 | 128 | 629 | 17 145 | 14.217 473 | 3.577 261 e+00 | 1.892 944 |
| 4 | 256 | 2196 | 94 766 | 12.307 520 | 1.909 953 e+00 | 1.872 958 |
| 4 | 512 | 7054 | 463 726 | 11.362 652 | 9.448 682 e-01 | 2.021 396 |
| 4 | 1024 | 21 128 | 2 076 876 | 11.049 173 | 3.134 787 e-01 | 3.014 138 |
| 4 | 2048 | 60 356 | 8 849 356 | 11.006 331 | 4.284 219 e-02 | 7.317 056 |
| 4 | 4096 | 166 029 | 36 499 567 | 11.001 824 | 4.507 368 e-03 | 9.504 924 |
| 6 | 256 | 18 | 98 | 69.074 698 | | |
| 6 | 512 | 162 | 2366 | 51.915 141 | 1.715 956 e+01 | |
| 6 | 1024 | 986 | 26 812 | 41.730 666 | 1.018 448 e+01 | 1.684 874 |
| 6 | 2048 | 4759 | 208 969 | 34.991 830 | 6.738 836 e+00 | 1.511 311 |
| 6 | 4096 | 19 895 | 1 306 723 | 30.765 057 | 4.226 774 e+00 | 1.594 321 |
| 6 | 8192 | 75 032 | 7 064 310 | 28.401 466 | 2.363 591 e+00 | 1.788 285 |
| 6 | 16 384 | 261 996 | 34 544 564 | 27.209 538 | 1.191 927 e+00 | 1.982 999 |
| 8 | 4096 | 8 | 32 | 141.880 913 | | |
| 8 | 8192 | 170 | 2394 | 105.253 777 | 3.662 714 e+01 | |
| 8 | 16 384 | 1525 | 43 929 | 85.929 387 | 1.932 439 e+01 | 1.895 384 |
| 8 | 32 768 | 9832 | 472 538 | 73.408 035 | 1.252 135 e+01 | 1.543 315 |
| 8 | 65 536 | 51 771 | 3 747 977 | 64.802 203 | 8.605 833 e+00 | 1.454 984 |
| 8 | 131 072 | 237 027 | 24 334 123 | 58.667 735 | 6.134 468 e+00 | 1.402 865 |

measured rates for sufficiently large values of $K$. Nevertheless we observe that quite large values of $K$ are needed to obtain a decent accuracy in $E$ and that the value of $K$ necessary to reach a fixed accuracy grows fast with the number $N$ of electrons. For larger numbers of particles we are not able to reach the asymptotics.

We also see that the choice $T = 0.25$ results in a substantially smaller problem size than the choice $T = 0$. For example, for $S = N/2$, $N = 4$ and $K = 2048$ we need to handle 448 481 degrees of freedom and 89 628 377 non-zero matrix entries for $T = 0$ but only 109 183 degrees of freedom and 20 867 911 non-zero matrix entries

TABLE 2. $d = 1$, $S = N/2$, $a = 20$, $D = 10$, $T = 0$, $Z = N$.

| $N$ | $K$ | $M_{\mathcal{A}^{(N,S)}}$ | $\#A$ | $E$ | $\Delta E$ | $\epsilon$ |
|---|---|---|---|---|---|---|
| 2 | 2 | 9 | 37 | 6.732 775 | | |
| 2 | 4 | 21 | 137 | 4.732 600 | 2.000 175 e+00 | |
| 2 | 8 | 57 | 605 | 3.011 013 | 1.721 588 e+00 | 1.161 820 |
| 2 | 16 | 141 | 2413 | 2.281 144 | 7.298 688 e-01 | 2.358 763 |
| 2 | 32 | 361 | 10 121 | 1.942 338 | 3.388 057 e-01 | 2.154 240 |
| 2 | 64 | 877 | 40 613 | 1.869 510 | 7.282 840 e-02 | 4.652 109 |
| 2 | 128 | 2081 | 161 853 | 1.864 274 | 5.235 401 e-03 | 13.910 760 |
| 2 | 256 | 4845 | 641 605 | 1.863 963 | 3.116 076 e-04 | 16.801 261 |
| 2 | 512 | 11 101 | 2 543 457 | 1.863 928 | 3.461 836 e-05 | 9.001 223 |
| 4 | 8 | 28 | 176 | 28.221 791 | | |
| 4 | 16 | 121 | 1425 | 20.497 004 | 7.724 787 e+00 | |
| 4 | 32 | 513 | 10 137 | 15.324 614 | 5.172 390 e+00 | 1.493 466 |
| 4 | 64 | 1869 | 56 249 | 11.844 463 | 3.480 151 e+00 | 1.486 255 |
| 4 | 128 | 6244 | 277 172 | 9.724 198 | 2.120 265 e+00 | 1.641 375 |
| 4 | 256 | 19 499 | 1 256 803 | 8.427 853 | 1.296 344 e+00 | 1.635 573 |
| 4 | 512 | 57 887 | 5 406 447 | 7.737 085 | 6.907 689 e-01 | 1.876 669 |
| 4 | 1024 | 163 796 | 22 247 576 | 7.444 972 | 2.921 127 e-01 | 2.364 734 |
| 4 | 2048 | 448 481 | 89 628 377 | 7.370 270 | 7.470 152 e-02 | 3.910 399 |
| 6 | 16 | 1 | 1 | 97.091 697 | | |
| 6 | 32 | 17 | 73 | 71.819 258 | 2.527 244 e+01 | |
| 6 | 64 | 181 | 2177 | 51.615 444 | 2.020 381 e+01 | 1.250 875 |
| 6 | 128 | 1148 | 24 788 | 39.657 137 | 1.195 831 e+01 | 1.689 521 |
| 6 | 256 | 5755 | 194 619 | 31.978 615 | 7.678 523 e+00 | 1.557 371 |
| 6 | 512 | 24 763 | 1 219 527 | 26.836 060 | 5.142 555 e+00 | 1.493 134 |
| 6 | 1024 | 96 044 | 6 601 008 | 23.302 798 | 3.533 262 e+00 | 1.455 469 |
| 6 | 2048 | 344 197 | 32 281 209 | 20.855 533 | 2.447 265 e+00 | 1.443 759 |
| 8 | 256 | 32 | 200 | 131.326 374 | | |
| 8 | 512 | 468 | 7316 | 99.493 684 | 3.183 269 e+01 | |
| 8 | 1024 | 3776 | 104 800 | 79.092 584 | 2.040 110 e+01 | 1.560 342 |
| 8 | 2048 | 23 113 | 992 209 | 65.080 337 | 1.401 225 e+01 | 1.455 948 |
| 8 | 4096 | 118 906 | 7 304 218 | 55.348 934 | 9.731 403 e+00 | 1.439 900 |

for $T = 0.25$ whereas the achieved accuracy is just slightly lower, *i.e.* we get the energy 7.767 587 instead of 7.370 270. This demonstrates the effect of the improved complexity of the generalized sparse grid in comparison to the regular sparse grid. Further experiments with varying values of $a$ showed similar results with respect to the convergence rates.

In Figures 4 and 5 we depict the zero/non-zero pattern of various stiffness matrices $A$ which gives an impression of their sparsity. Here, we show matrices in the ordering of our algorithm and in the so-called reverse Cuthill-McKee ordering [61] which results in matrices with reduced band widths.

Now we consider the case of two-dimensional particles. Again, we use the generalized antisymmetric sparse grid space $V_{K,T}^{\mathcal{A}^{(N,S)}}$ from (32) and focus on the two cases $T = 0$ and $T = 0.25$ with $S = \lfloor N/2 \rfloor$. Tables 5 and 6 give the obtained results. We see an analogous behaviour of the convergence rates and complexities as in the one-dimensional case. Now however it is even more difficult to achieve a decent accuracy and, already for a quite small number of electrons $N > 2$, we are not able to reach the asymptotics due to the huge size of the matrices involved.

TABLE 3. $d = 1$, $S = 0$, $a = 20$, $D = 10$, $T = 0.25$, $Z = N$.

| $N$ | $K$ | $M_{\mathcal{A}^{(N,S)}}$ | $\#A$ | $E$ | $\Delta E$ | $\epsilon$ |
|---|---|---|---|---|---|---|
| 2 | 2 | 4 | 12 | 9.477 604 | | |
| 2 | 4 | 8 | 40 | 7.672 409 | 1.805 195 e+00 | |
| 2 | 8 | 21 | 177 | 5.203 866 | 2.468 543 e+00 | 0.731 280 |
| 2 | 16 | 50 | 740 | 3.923 744 | 1.280 121 e+00 | 1.928 367 |
| 2 | 32 | 131 | 3311 | 3.006 528 | 9.172 167 e-01 | 1.395 658 |
| 2 | 64 | 304 | 13 268 | 2.797 567 | 2.089 609 e-01 | 4.389 419 |
| 2 | 128 | 695 | 53 323 | 2.759 665 | 3.790 167 e-02 | 5.513 236 |
| 2 | 256 | 1553 | 212 021 | 2.758 630 | 1.035 467 e-03 | 36.603 456 |
| 2 | 512 | 3409 | 843 669 | 2.758 536 | 9.356 699 e-05 | 11.066 585 |
| 4 | 32 | 8 | 32 | 32.074 608 | | |
| 4 | 64 | 36 | 316 | 24.541 347 | 7.533 261 e+00 | |
| 4 | 128 | 151 | 2503 | 19.247 385 | 5.293 962 e+00 | 1.422 991 |
| 4 | 256 | 532 | 15 002 | 15.465 571 | 3.781 814 e+00 | 1.399 847 |
| 4 | 512 | 1631 | 73 357 | 13.247 843 | 2.217 728 e+00 | 1.705 265 |
| 4 | 1024 | 4622 | 324 740 | 12.026 948 | 1.220 895 e+00 | 1.816 477 |
| 4 | 2048 | 12 462 | 1 357 806 | 11.353 865 | 6.730 834 e-01 | 1.813 884 |
| 4 | 4096 | 32 114 | 5 499 708 | 11.072 627 | 2.812 379 e-01 | 2.393 288 |
| 4 | 8192 | 79 498 | 21 813 126 | 11.011 562 | 6.106 479 e-02 | 4.605 566 |
| 6 | 1024 | 20 | 116 | 68.871 561 | | |
| 6 | 2048 | 144 | 2104 | 53.571 662 | 1.529 990 e+01 | |
| 6 | 4096 | 739 | 19 571 | 44.680 401 | 8.891 261 e+00 | 1.720 779 |
| 6 | 8192 | 2981 | 124 625 | 38.182 233 | 6.498 168 e+00 | 1.368 272 |
| 6 | 16 384 | 10 570 | 663 440 | 33.724 487 | 4.457 747 e+00 | 1.457 725 |
| 6 | 32 768 | 34 456 | 3 152 064 | 30.603 389 | 3.121 097 e+00 | 1.428 263 |
| 6 | 65 536 | 104 068 | 13 725 564 | 28.670 832 | 1.932 558 e+00 | 1.615 009 |
| 6 | 131 072 | 297 605 | 56 837 725 | 27.571 226 | 1.099 606 e+00 | 1.757 500 |
| 8 | 32 768 | 6 | 22 | 144.732 166 | | |
| 8 | 65 536 | 84 | 928 | 114.329 594 | 3.040 257 e+01 | |
| 8 | 131 072 | 648 | 14 888 | 94.153 265 | 2.017 633 e+01 | 1.506 844 |
| 8 | 262 144 | 3457 | 133 133 | 81.977 181 | 1.217 608 e+01 | 1.657 046 |
| 8 | 524 288 | 15 421 | 903 461 | 72.775 532 | 9.201649 e+00 | 1.323 250 |
| 8 | 1 048 576 | 60 576 | 5 090 202 | 65.986 443 | 6.789 090 e+00 | 1.355 358 |
| 8 | 2 097 152 | 215 864 | 25 293 976 | 60.838 970 | 5.147 473 e+00 | 1.318 917 |

Finally we consider the case of three-dimensional electrons. Again, we use the generalized antisymmetric sparse grid space $V_{K,T}^{\mathcal{A}^{(N,S)}}$ from (32) and focus on the two cases $T = 0$ and $T = 0.25$ with either $S = 0$ or $S = \lfloor N/2 \rfloor$. Tables 7–10 give the obtained results. Due to the associated complexity we are presently only able to treat hydrogen and helium, *i.e.* the case $N = 1$ and $N = 2$. Also the resolution $K$ is very low. This is due to the fact that now, in the case $d = 3$, the number of degrees of freedom $M_{\mathcal{A}^{(N,S)}}$ scales roughly like $K^3$ albeit with a large proportionality constant and the amount $\#A$ of non-zero-entries also grows, not quite like $M^2_{\mathcal{A}^{(N,S)}}$ due to the zero-entries, but nevertheless fast. Thus we are only able to deal with a resolution up to $K = 16$ for $N = 2$, which is by far not sufficient to get into any asymptotics.

Nevertheless, from our tables we can see some convergence. For hydrogen we come within two digits to the exact value $E = -0.5$. For helium we only can compare with values from the literature. In [70] we find the value $-2.9037$ for the non-relativistic energy of the 1S state of neutral helium. So our computed results tend at least in the right direction. Since we use a grid based approach we are allowed to use extrapolation techniques for

TABLE 4. $d = 1$, $S = N/2$, $a = 20$, $D = 10$, $T = 0.25$, $Z = N$.

| $N$ | $K$ | $M_{\mathcal{A}^{(N,S)}}$ | $\#A$ | $E$ | $\Delta E$ | $\epsilon$ |
|---|---|---|---|---|---|---|
| 2 | 2 | 9 | 37 | 6.732 775 | | |
| 2 | 4 | 17 | 105 | 5.952 855 | 7.799 202 e-01 | |
| 2 | 8 | 45 | 449 | 3.641 773 | 2.311 081 e+00 | 0.337 470 |
| 2 | 16 | 105 | 1717 | 2.744 973 | 8.968 008 e-01 | 2.577 029 |
| 2 | 32 | 269 | 7285 | 2.093 947 | 6.510 254 e-01 | 1.377 521 |
| 2 | 64 | 617 | 28 185 | 1.916 137 | 1.778 098 e-01 | 3.661 358 |
| 2 | 128 | 1405 | 110 901 | 1.868 440 | 4.769 721 e-02 | 3.727 887 |
| 2 | 256 | 3125 | 434 441 | 1.864 333 | 4.107 707 e-03 | 11.611 639 |
| 4 | 8 | 12 | 52 | 34.786 234 | | |
| 4 | 16 | 44 | 344 | 27.914 580 | 6.871 653 e+00 | |
| 4 | 32 | 225 | 3425 | 18.856 317 | 9.058 263 e+00 | 0.758 606 |
| 4 | 64 | 681 | 16 049 | 15.299 227 | 3.557 090 e+00 | 2.546 537 |
| 4 | 128 | 2083 | 74 591 | 12.651 914 | 2.647 313 e+00 | 1.343 660 |
| 4 | 256 | 6066 | 324 330 | 10.551 109 | 2.100 805 e+00 | 1.260 143 |
| 4 | 512 | 16 548 | 1 332 468 | 9.123 509 | 1.427 600 e+00 | 1.471 564 |
| 4 | 1024 | 43 147 | 5 306 371 | 8.296 504 | 8.270 051 e-01 | 1.726 228 |
| 4 | 2048 | 109 183 | 20 867 911 | 7.767 587 | 5.289 171 e-01 | 1.563 582 |
| 6 | 32 | 1 | 1 | 97.091 697 | | |
| 6 | 64 | 17 | 73 | 71.819 258 | 2.527 244 e+01 | |
| 6 | 128 | 111 | 1059 | 56.350 524 | 1.546 873 e+01 | 1.633 776 |
| 6 | 256 | 594 | 10 782 | 45.233 905 | 1.111 662 e+01 | 1.391 496 |
| 6 | 512 | 2624 | 74 272 | 36.638 004 | 8.595 900 e+00 | 1.293 247 |
| 6 | 1024 | 9407 | 391 655 | 31.515 569 | 5.122 435 e+00 | 1.678 089 |
| 6 | 2048 | 31 923 | 1 895 923 | 27.576 464 | 3.939 106 e+00 | 1.300 406 |
| 6 | 4096 | 101 536 | 8 454 432 | 24.426 415 | 3.150 049 e+00 | 1.250 490 |
| 6 | 8192 | 299 917 | 34 951 561 | 22.048 084 | 2.378 330 e+00 | 1.324 479 |
| 8 | 1024 | 32 | 200 | 131.326 374 | | |
| 8 | 2048 | 280 | 3744 | 107.195 449 | 2.413 093 e+01 | |
| 8 | 4096 | 2016 | 48 724 | 86.966 479 | 2.022 897 e+01 | 1.192 889 |
| 8 | 8192 | 10 096 | 369 224 | 73.127 731 | 1.383 875 e+01 | 1.461763 |
| 8 | 16 384 | 42 693 | 2 226 613 | 63.481 008 | 9.646 723 e+00 | 1.434 554 |
| 8 | 32 768 | 164 061 | 11 775 961 | 56.129 185 | 7.351 824 e+00 | 1.312 154 |

TABLE 5. $d = 2$, $S = N/2$, $a = 20$, $D = 10$, $T = 0$, $Z = N$.

| $N$ | $K$ | $M_{\mathcal{A}^{(N,S)}}$ | $\#A$ | $E$ | $\Delta E$ | $\epsilon$ |
|---|---|---|---|---|---|---|
| 2 | 2 | 49 | 1297 | 1.312 598 | | |
| 2 | 4 | 225 | 14 889 | 0.170 526 | 1.142 072 e+00 | |
| 2 | 8 | 1537 | 244 601 | –0.836 206 | 1.006 732 e+00 | 1.134 435 |
| 2 | 16 | 8321 | 3 759 217 | –1.177 562 | 3.413 557 e-01 | 2.949 216 |
| 2 | 32 | 49 921 | 64 062 257 | –1.271 094 | 9.353 198e-02 | 3.649 615 |
| 4 | 4 | 64 | 1112 | 8.302 101 | | |
| 4 | 8 | 1408 | 85 048 | 4.310 960 | 3.991 141 e+00 | |
| 4 | 16 | 14 736 | 2 102 944 | 0.450 265 | 3.860 696 e+00 | 1.033 788 |
| 4 | 32 | 166 864 | 54 529 984 | –3.439 995 | 3.890 259 e+00 | 0.992 401 |

TABLE 6. $d = 2$, $S = N/2$, $a = 20$, $D = 10$, $T = 0.25$, $Z = N$.

| $N$ | $K$ | $M_{\mathcal{A}^{(N,S)}}$ | $\#A$ | $E$ | $\Delta E$ | $\epsilon$ |
|---|---|---|---|---|---|---|
| 2 | 2 | 49 | 1297 | 1.312 598 | | |
| 2 | 4 | 161 | 13 281 | 1.172 373 | 1.402 244 e-01 | |
| 2 | 8 | 897 | 183 209 | –0.514 416 | 1.686 790 e+00 | 0.083 131 |
| 2 | 16 | 4289 | 2 679 929 | –0.920 735 | 4.063 190 e-01 | 4.151 393 |
| 2 | 32 | 25 345 | 44 126 289 | –1.236 727 | 3.159 916 e-01 | 1.285 854 |
| 4 | 8 | 320 | 13 048 | 5.945 370 | | |
| 4 | 16 | 2560 | 304 760 | 4.138 734 | 1.806 635 e+00 | |
| 4 | 32 | 37 392 | 10 748 704 | –1.004 031 | 5.142 766 e+00 | 0.351 296 |

TABLE 7. $d = 3$, $S = 0$, $a = 15$, $D = 7.5$, $T = 0$, $Z = N$.

| $N$ | $K$ | $M_{\mathcal{A}^{(N,S)}}$ | $\#A$ | $E$ | $\Delta E$ | $\epsilon$ |
|---|---|---|---|---|---|---|
| 1 | 2 | 125 | 14 953 | –0.360 205 | | |
| 1 | 4 | 729 | 518 697 | –0.441 314 | 8.110 947 e-02 | |
| 1 | 8 | 4913 | 2 384 2801 | –0.485 015 | 4.370 059 e-02 | 1.856 027 |
| 1 | 16 | 35 937 | 1 283 473 497 | –0.497 398 | 1.238 257 e-02 | 3.529 201 |
| 2 | 2 | 124 | 14 716 | –1.028 674 | | |
| 2 | 4 | 1053 | 534 549 | –1.445 113 | 4.164 390 e-01 | |
| 2 | 8 | 18 206 | 28 670 240 | –1.839 117 | 3.940 037 e-01 | 1.056 942 |
| 2 | 16 | 210 907 | 1 669 143 059 | –2.055 057 | 2.159 402 e-01 | 1.824 596 |

TABLE 8. $d = 3$, $S = N/2$, $a = 15$, $D = 7.5$, $T = 0$, $Z = N$.

| $N$ | $K$ | $M_{\mathcal{A}^{(N,S)}}$ | $\#A$ | $E$ | $\Delta E$ | $\epsilon$ |
|---|---|---|---|---|---|---|
| 2 | 2 | 249 | 30 141 | –1.084 312 | | |
| 2 | 4 | 2133 | 1 080 285 | –1.491 861 | 4.075 485 e-01 | |
| 2 | 8 | 36 537 | 58 141 181 | –1.903 633 | 4.117 722 e-01 | 0.989 743 |

TABLE 9. $d = 3$, $S = 0$, $a = 15$, $D = 7.5$, $T = 0.25$, $Z = N$.

| $N$ | $K$ | $M_{\mathcal{A}^{(N,S)}}$ | $\#A$ | $E$ | $\Delta E$ | $\epsilon$ |
|---|---|---|---|---|---|---|
| 2 | 2 | 124 | 14 716 | –1.028 674 | | |
| 2 | 4 | 728 | 517 312 | –1.443 261 | 4.145 863 e-01 | |
| 2 | 8 | 7785 | 24 311 681 | –1.834 115 | 3.908 549 e-01 | 1.060 717 |
| 2 | 16 | 74 918 | 1 331 689 684 | –2.044 871 | 2.107 553 e-01 | 1.854 543 |

the error. We then obtain with the data from Tables 7 and 10 the extrapolated values $-0.49\,956$ and $-2.45\,407$, respectively.

Altogether we clearly see that we do not reach the asymptotics with respect to $K$. The Fourier basis simply does not decay fast enough to represent the Kato-cusps of the solution properly with moderate values of $K$. Note that, in particular in a sparse grid approach, the values for the discretization parameter $K$ have to be chosen quite high to resolve the neighborhood of an electron-electron cusp at the diagonal sufficiently.

To treat practically relevant problems with $d = 3$ we thus learned that we have to use better basis functions than the regular Fourier system. This is due to the globality of the Fourier functions which are not able to locally resolve singularities like the nuclei-electron cusps and the electron-electron cusps. Therefore we need to employ more suited function systems here. A promising approach might be to introduce an additional mapping

TABLE 10. $d = 3$, $S = N/2$, $a = 15$, $D = 7.5$, $T = 0.25$, $Z = N$.

| $N$ | $K$ | $M_{\mathcal{A}^{(N,S)}}$ | $\#A$ | $E$ | $\Delta E$ | $\epsilon$ |
|---|---|---|---|---|---|---|
| 2 | 4 | 1457 | 1 038 850 | –1.476 170 | | |
| 2 | 8 | 15 597 | 48 605 034 | –1.871 390 | 3.952 200 e-01 | |
| 2 | 16 | 149 961 | 2 665 724 933 | –2.121 319 | 2.499 289 e-01 | 1.581 330 |

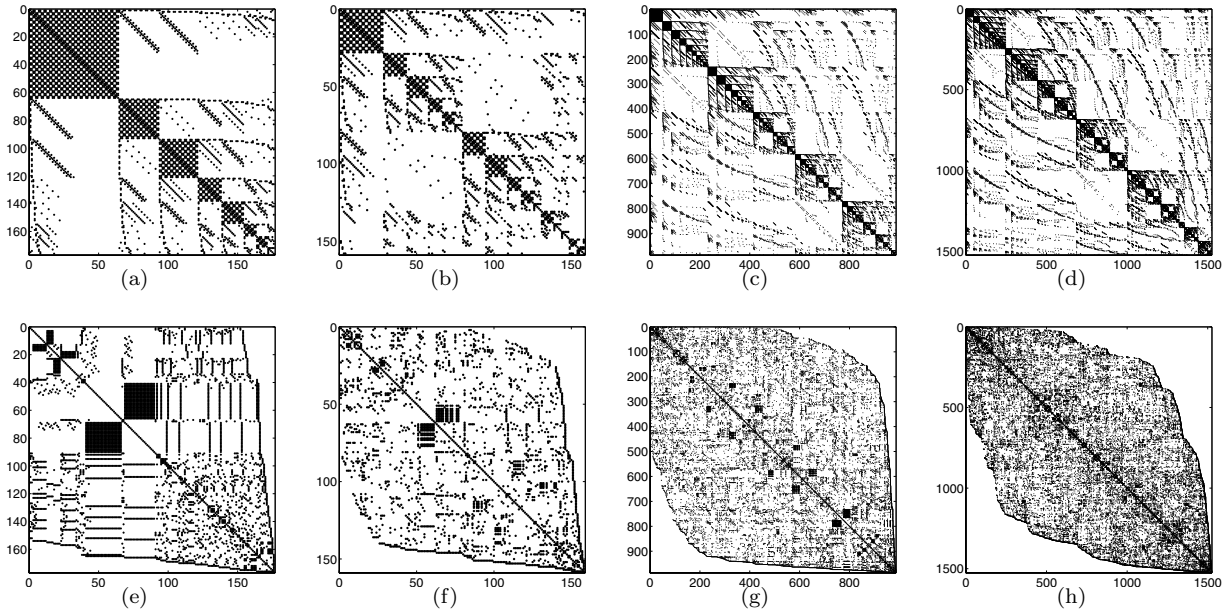


FIGURE 4. Non-zero entries of matrices $A$ for $d = 1$, $a = 20$, $D = 10$, $T = 0$, $Z = N$ and $S = 0$. (a), (e): $N = 2$, $K = 32$. (b), (f): $N = 4$, $K = 64$. (c), (g): $N = 6$, $K = 1024$. (d), (h): $N = 8$, $K = 16\,384$. The matrices (a)–(d) are depicted in the ordering of our algorithm and the matrices (e)–(h) are depicted in the reverse Cuthill-McKee ordering.

which resembles a transformation of the coordinate system and allows for a grading of the mesh in $\vec{\mathbf{k}}$-space towards the cusps, see for example [34] for a first attempt in this direction involving sparse grids. Such a mapping of the Fourier grid is related to the introduction of a Riemannian metric into the problem, which even can be done in an adaptive fashion, see [1, 24, 44–46] for details. Another approach might be the use of localized multiscale basis functions like interpolets, prewavelets or wavelets which preferably additionally fulfil the orthonormality condition and moreover allow for local adaptivity. Such function systems can then be employed in our generalized sparse grid approach and should result in substantially better approximations already in the (practically reachable) preasymptotic case. This however is future work. For first results in this direction, see [28]. However, another reason why we might not reach the asymptotics with respect to $K$ might be the term $\|\psi^{\mathcal{A}^{(N,S)}}\|_{\mathcal{H}^{t,l}_{\mathrm{mix}}}$ in the error estimates for antisymmetric generalized sparse grids. This term nevertheless might grow exponentially with the number $N$ of electrons and thus postpones the onset of convergence. This is a principal problem with the sparse grid approach.

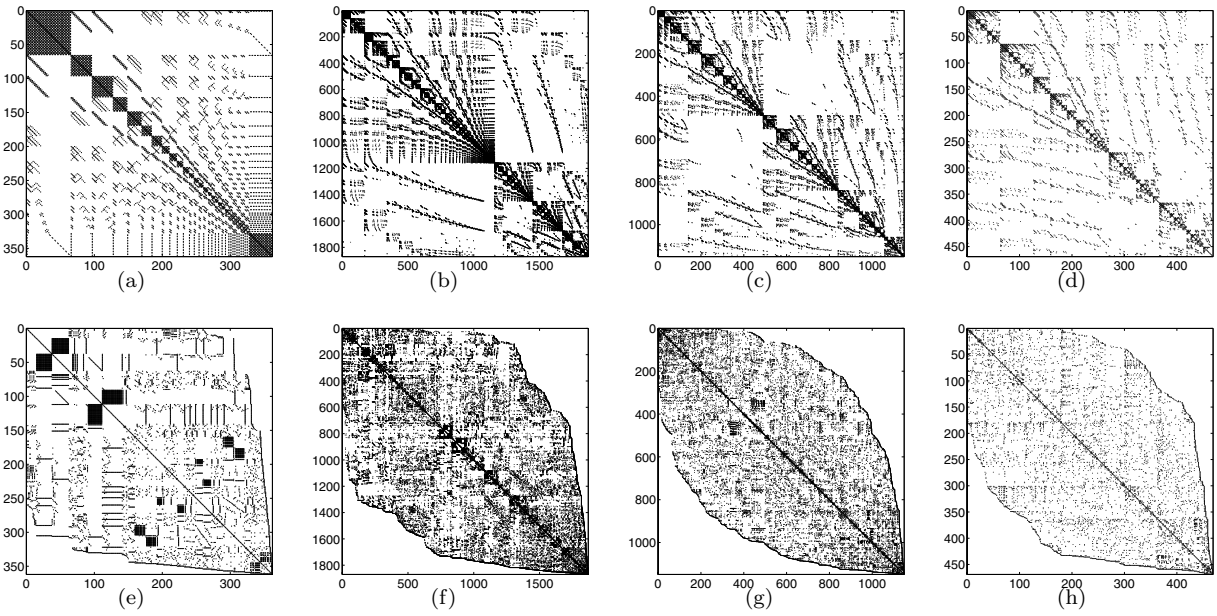


FIGURE 5. Non-zero entries of matrices $A$ for $d = 1$, $a = 20$, $D = 10$, $T = 0$, $Z = N$ and $S = N/2$. (a), (e): $N = 2$, $K = 32$. (b), (f): $N = 4$, $K = 64$. (c), (g): $N = 6$, $K = 128$. (d), (h): $N = 8$, $K = 512$. The matrices (a)–(d) are depicted in the ordering of our algorithm and the matrices (e)–(h) are depicted in the reverse Cuthill-McKee ordering.

## 6. Concluding remarks

In this article we discussed the sparse grid approach for the electronic Schrödinger equation. Here, we employed for the $d$-dimensional one-particle space the anisotropic product of a one-dimensional multiscale basis. A further product approach then gives a multiscale basis for the $N$-particle system. Truncation leads to different variants of sparse grid subspaces. Here, besides the conventional sparse grid approach we focused on optimized sparse grids which allow to take advantage of certain mixed smoothness properties of the function to be represented. We discussed the associated complexities and approximation properties. Then we generalized the sparse grid approach to the case of antisymmetry. To this end, the conventional product was replaced by the outer product which involves the Slater determinant construction. Additional conditions on the level indices of the multivariate basis were imposed which reflect the Pauli principle. We thus obtained a true basis for antisymmetric sparse grid spaces with a substantially reduced amount of degree of freedoms and derived the associated complexities and approximation properties. Then, we applied the Galerkin approach for the electronic Schrödinger equation using our antisymmetric sparse grid spaces. We set up the stiffness matrix and discussed its non-zero structure which results from the Slater-Condon rules. Furthermore we solve the associated discrete eigenvalue problem with a Lanczos solver. Finally, we applied our approach to model problems and compared costs, accuracy, convergence rate and scalability with respect to the number of electrons present in the system.

For reasons of simplicity we employed the Fourier basis as one-particle functions. Note that our approach is by no means restricted to this specific choice of multilevel basis. Any multilevel basis for the one-particle space with a sufficient decay property may be used as basic ingredient for our sparse grid approach with similar results. Candidates are other hierarchical global polynomial systems or function families with localization properties like wavelets, interpolets, multiscale finite element systems and related frames or multiscale Gaussians. Of course, if the functions are no longer orthogonal the resulting system matrix assembly may be dense and is thus in general

more costly. Note that a wavelet-like system with localization properties might further improve our complexity results due to a possibly adaptive local resolution of nuclei-electron cusps and electron-electron cusps.

In any case we learned that in principle the sparse grid approach possesses favourable approximation rates and cost complexities which in the case $0 < T < 1$ exhibit no exponential dependency of the number $N$ of particles with respect to the discretization parameter $K$. Note however that the involved order constants still depend on $N$. Moreover, since in our (not yet adaptive) approach at least the one- and in particular the two-electron interactions in the discrete solution are resolved with a uniform grid which involves $O(K^6)$ degrees of freedom for the case $d = 3$, our computations are still limited due to this huge (but to some extent constant with respect to $N$) number of degrees of freedoms and associated operations and we by far do not reach the asymptotics. We nevertheless believe that for larger $N$ the sparse grid effect in principle kicks in, *i.e.* that mainly pair interactions must be resolved properly but that triple and higher interactions are greatly sparsified. To this end better one- and two-particle basis function sets are needed within our sparse grid approach in the future. The term $\|\psi^{\mathcal{A}^{(N,S)}}\|_{\mathcal{H}^{t,l}_{\mathrm{mix}}}$ in the error estimates nevertheless might grow exponentially with the number $N$ of electrons and thus postpones the onset of convergence.

*Acknowledgements.* This work was partially supported by the Deutsche Forschungsgemeinschaft in the priority program 1145 "Modern and universal first-principles methods for many-electron systems in chemistry and physics". We are grateful to Harry Yserentant (TU Berlin) and Martin Mohlenkamp (U. Ohio) for fruitful discussions and suggestions.

## References

[1] E. Ackad and M. Horbatsch, Numerical solution of the Dirac equation by a mapped Fourier grid method. *J. Phys. A: Math. General* **38** (2005) 3157–3171.
[2] R.A. Adams, *Sobolev spaces*. Academic Press, New York (1975).
[3] A. Arnold, Mathematical concepts of open quantum boundary conditions. *Transport Theory Statist. Phys.* **30** (2001) 561–584.
[4] P.W. Atkins and R.S. Friedman, *Molecular quantum mechanics*. Oxford University Press, Oxford (1997).
[5] K.I. Babenko, Approximation by trigonometric polynomials in a certain class of periodic functions of several variables. *Dokl. Akad. Nauk SSSR* **132** (1960) 672–675.
[6] S. Balay, K. Buschelman, V. Eijkhout, W.D. Gropp, D. Kaushik, M.G. Knepley, L.C. McInnes, B.F. Smith and H. Zhang, *PETSc users manual*. Tech. Report ANL-95/11 - Revision 2.1.5, Argonne National Laboratory (2004).
[7] R. Bellmann, *Adaptive control processes: A guided tour*. Princeton University Press (1961).
[8] J. Boyd, *Chebyshev and Fourier spectral methods*. Dover Publications, New York (2000).
[9] H. Bungartz, *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung*. Dissertation, Institut für Informatik, TU München (1992).
[10] H. Bungartz, *Finite elements of higher order on sparse grids*. Habilitationsschrift, Institut für Informatik, TU München and Shaker Verlag, Aachen (1998).
[11] H. Bungartz and M. Griebel, A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives. *J. Complexity* **15** (1999) 167–199.
[12] H. Bungartz and M. Griebel, Sparse grids. *Acta Numer.* **13** (2004) 147–269.
[13] Z. Cai, J. Mandel and S. McCormick, Multigrid methods for nearly singular linear equations and eigenvalue problems. *SIAM J. Numer. Anal.* **34** (1997) 178–200.
[14] T. Chan and I. Sharapov, Subspace correction multi-level methods for elliptic eigenvalue problems. *Numer. Linear Algebra Appl.* **9** (2002) 1–20.
[15] C. Chui and Y. Wang, A general framework for compactly supported splines and wavelets. *J. Approx. Theory* **71** (1992) 263–304.
[16] A. Cohen, *Numerical analysis of wavelet methods, Studies in Mathematics and its Applications* **32**. North Holland (2003).
[17] E. Condon, The theory of complex spectra. *Phys. Rev.* **36** (1930) 1121–1133.
[18] I. Daubechies, *Ten lectures on wavelets*. CBMS-NSF Regional Conf. Series in Appl. Math. **61**, SIAM (1992).
[19] G. Deslauriers and S. Dubuc, Symmetric iterative interpolation processes. *Constr. Approx.* **5** (1989) 49–68.
[20] R. DeVore, S. Konyagin and V. Temlyakov, Hyperbolic wavelet approximation. *Constr. Approx.* **14** (1998) 1–26.
[21] N. Dobrovol'skii and A. Roshchenya, Number of lattice points in the hyperbolic cross. *Math. Notes* **11** (1998) 319–324.
[22] D. Donoho and P. Yu, *Deslauriers-Dubuc: Ten years after*, *CRM Proceedings and Lecture Notes* **18**, G. Deslauriers and S. Dubuc Eds. (1999).
[23] P. Ewald, Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* **64** (1921) 253–287.

[24] E. Fattal, R. Baer and R. Kosloff, Phase space approach for optimizing grid representations: the mapped Fourier method. *Phys. Rev. E* **53** (1996) 1217–1227.
[25] T. Fevens and H. Jiang, Absorbing boundary conditions for the Schrödinger equation. *SIAM J. Scientific Comput.* **21** (1999) 255–282.
[26] R. Feynman, *There's plenty of room at the bottom: An invitation to enter a new world of physics*. Engineering and Science XXIII, Feb. issue (1960), `http://www.zyvex.com/nanotech/feynman.html`.
[27] H.-J. Flad, W. Hackbusch, D. Kolb and R. Schneider, Wavelet approximation of correlated wavefunctions. I. Basics., *J. Chem. Phys.* **116** (2002) 9641–9857.
[28] H.-J. Flad, W. Hackbusch and R. Schneider, *Best N term approximation in electronic structure calculations. I. One electron reduced density matrix*. Tech. Report 05-9, Berichtsreihe des Mathematischen Seminars der Universität Kiel (2005).
[29] H. Fliegl, W. Klopper and C. Hättig, Coupled-cluster theory with simplified linear-r12 corrections: The CCSD(R12) model. *J. Chem. Phys.* **122** (2005) 084107.
[30] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Sørensen, The electron density is smooth away from the nuclei. *Commun. Math. Phys.* **228** (2002) 401–415.
[31] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Sørensen, Sharp regularity results for Coulombic many-electron wave functions. *Commun. Math. Phys.* **255** (2005) 183–227.
[32] K. Frank, S. Heinrich and S. Pereverzev, Information complexity of multivariate Fredholm equations in Sobolev classes. *J. Complexity* **12** (1996) 17–34.
[33] G. Friesecke, *The configuraton-interaction equations for atoms and molecules: Charge quantization and existence of solutions*. Preprint, June 28, 1999, Mathematical Insitute, University of Oxford, UK.
[34] 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. Comput. Phys.* **165** (2000) 694–716.
[35] T. Gerstner and M. Griebel, Numerical integration using sparse grids. *Numer. Algorithms* **18** (1998) 209–232.
[36] T. Gerstner and M. Griebel, Dimension-adaptive tensor-product quadrature. *Computing* **71** (2003) 65–87.
[37] M. Griebel, Multilevel algorithms considered as iterative methods on semidefinite systems. *SIAM J. Sci. Stat. Comput.* **15** (1994) 547–565.
[38] M. Griebel, Sparse grids and related approximation schemes for higher dimensional problems, in *Proceedings of the conference on Foundations of Computational Mathematics (FoCM05)*, Santander, Spain, 2005.
[39] M. Griebel and S. Knapek, Optimized tensor-product approximation spaces. *Constr. Approx.* **16** (2000) 525–540.
[40] M. Griebel and P. Oswald, On additive Schwarz preconditioners for sparse grid discretizations. *Numer. Math.* **66** (1994) 449–463.
[41] M. Griebel and P. Oswald, On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.* **70** (1995) 161–180.
[42] M. Griebel and P. Oswald, Tensor product type subspace splitting and multilevel iterative methods for anisotropic problems. *Adv. Comput. Math.* **4** (1995) 171–206.
[43] M. Griebel, P. Oswald and T. Schiekofer, Sparse grids for boundary integral equations. *Numer. Mathematik* **83** (1999) 279–312.
[44] F. Gygi, Adaptive Riemannian metric for plane-wave electronic-structure calculations. *Europhys. Lett.* **19** (1992) 617.
[45] F. Gygi, Electronic-structure calculations in adaptive coordinates. *Phys. Rev. B* **48** (1993) 11692.
[46] D. Hamann, Comparison of global and local adaptive coordinates for density-functional calculations. *Phys. Rev. B* **63** (2001) 075107.
[47] V. Hernandez, J. Roman and V. Vidal, SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems. *ACM Trans. Math. Software* **31** (2005) 351–362.
[48] R. Hochmuth, *Wavelet bases in numerical analysis and restricted nonlinear approximation*. Habilitationsschrift, Freie Universität Berlin (1999).
[49] R. Hochmuth, S. Knapek and G. Zumbusch, *Tensor products of Sobolev spaces and applications*. Tech. Report 685, SFB 256, Univ. Bonn (2000).
[50] M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Sørensen, Electron wavefunction and densities for atoms. *Ann. Henri Poincaré* **2** (2001) 77–100.
[51] G. Karniadakis and S. Sherwin, *Spectral/hp element methods for CFD*. Oxford University Press (1999).
[52] T. Kato, On the eigenfunctions of many-particle systems in quantum mechanics. *Commun. Pure Appl. Math.* **10** (1957) 151–177.
[53] J. Keller and D. Givoli, Exact non-reflecting boundary conditions. *J. Comput. Phys.* **82** (1989) 172–192.
[54] S. Knapek, *Approximation und Kompression mit Tensorprodukt-Multiskalenräumen*. Dissertation, Universität Bonn, April (2000).
[55] S. Knapek, *Hyperbolic cross approximation of integral operators with smooth kernel*. Tech. Report 665, SFB 256, Univ. Bonn (2000).
[56] A. Knyazev and K. Neymeyr, Efficient solution of symmetric eigenvalue problem using multigrid preconditioners in the locally optimal block conjugate gradient method. *Electronic Trans. Numer. Anal.* **15** (2003) 38–55.

[57] W. Kutzelnigg, *Convergence of expansions in a Gaussian basis*. Strategies and Applications in Quantum Chemistry, M. Defranceschi and Y. Ellinger Eds., Kluwer, Dordrecht (1996).
[58] W. Kutzelnigg and D. Mukherjee, Minimal parametrization of an n-electron state. *Phys. Rev. A* **71** (2005) 022502.
[59] C. Le Bris, Computational chemistry from the perspective of numerical analysis. *Acta Numer.* **14** (2005) 363–444.
[60] I. Levine, *Quantum chemistry*, 5th edn., Prentice-Hall (2000).
[61] W. Liu and A. Sherman, Comparative analysis of Cuthill-McKee and reverse Cuthill-McKee ordering algorithms for sparse matrices. *SIAM J. Numer Anal.* **13** (1976) 198–213.
[62] O. Livne and A. Brandt, $O(N \log N)$ *multilevel calculation of N eigenfunctions*. Multiscale Computational Methods in Chemistry and Physics, A. Brandt, J. Bernholc and K. Binder Eds., NATO Science Series III: Computer and Systems Sciences, IOS Press **177** (2001) 112–136.
[63] V. Maz'ya and G. Schmidt, On approximate approximations using Gaussian kernels. *IMA J. Numer. Anal.* **16** (1996) 13–29.
[64] D. Mazziotti, Variational two-electron reduced density matrix theory for many-electron atoms and molecules: Implementation of the spin- and symmetry-adapted T-2 condition through first-order semidefinite programming. *Phys. Rev. A* **72** (2005) 032510.
[65] A. Messiah, *Quantum mechanics*. Vol. 1 and 2, North-Holland, Amsterdam, 1961/62.
[66] P. Nitsche, *Best n term approximation spaces for sparse grids*. Tech. Report 2003-11, ETH Zürich, Seminar für Angewandte Mathematik (2003).
[67] P. Oswald, *Multilevel finite element approximation*. Teubner Skripten zur Numerik, Teubner, Stuttgart (1994).
[68] R. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules*. Oxford University Press, New York (1989).
[69] H.-J. Schmeisser and H. Triebel, *Fourier analysis and functions spaces*. John Wiley, Chichester (1987).
[70] J.S. Sims and S.A. Hagstrom, High-precision Hy-CI variational calculations for the ground state of neutral helium and helium-like ions. *Int. J. Quant. Chem.* **90** (2002) 1600–1609.
[71] J. Slater, The theory of complex spectra. *Phys. Rev.* **34** (1929) 1293–1322.
[72] S. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions. *Soviet Math. Dokl.* **4** (1963) 240–243, Russian original in *Dokl. Akad. Nauk SSSR* **148** (1963) 1042–1045.
[73] B. Szabo and I. Babuska, *Finite element analysis*. Wiley (1991).
[74] J. Szeftel, Design of absorbing boundary conditions for Schrödinger equations in $\mathbb{R}^{d*}$. *SIAM J. Numer. Anal.* **42** (2004) 1527–1551.
[75] J. Weidmann, *Linear operators in Hilbert spaces*. Springer, New York (1980).
[76] H. Yserentant, *On the electronic Schrödinger equation*. Report 191, SFB 382, Univ. Tübingen (2003).
[77] H. Yserentant, On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives. *Numer. Math.* **98** (2004) 731–759.
[78] D. Zung, The approximation of classes of periodic functions of many variables. *Russian Math. Surveys* **38** (1983) 117–118.
[79] D. Zung, Approximation by trigonometric polynomials of functions of several variables on the torus. *Math. USSR Sbornik* **59** (1988) 247–267.