A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives

Hans-Joachim Bungartz,
Institut für Informatik,
TU München, D-80290 München

Michael Griebel,
Institut für Angewandte Mathematik,
Universität Bonn, Wegelerstr. 6, D-53115 Bonn

Abstract

We study the complexity of solving the d-dimensional Poisson equation on $]0,1[^d$. We restrict ourselves to cases where the solution u lies in some space of functions of bounded mixed derivatives (with respect to the L_∞- or the L_2-norm) up to $\partial^{2d}/\prod_{j=1}^{d}\partial x_j^2$. An upper bound for the complexity of computing a solution of some prescribed accuracy ε with respect to the energy norm is given, which is proportional to ε^{-1}. We show this result in a constructive manner by proposing a finite element method in a special sparse grid space [Bu1], which is obtained by an a priori grid optimization process based on the energy norm.

Thus, the result of this paper is twofold: First, from a theoretical point of view concerning the complexity of solving elliptic PDEs, a strong tractability result of the order $O(\varepsilon^{-1})$ is given, and, second, we provide a practically usable hierarchical basis finite element method of this complexity $O(\varepsilon^{-1})$, i. e. without logarithmic terms growing exponentially in d, at least for our sparse grid setting with its underlying smoothness requirements.

Keywords: Poisson equation, complexity, sparse grids, finite element method, multilevel methods.

AMS(MOS) subject classification: 35J05, 65N15, 65N30, 65N55, 65Y20, 68Q25.

1 Introduction

In this paper, we want to discuss the ε-complexity [TWW] of the solution of the d-dimensional Poisson equation $-\Delta u = f$ on the d-dimensional unit cube. That is, for some class of functions F, the computational effort that is necessary to compute an approximation of accuracy ε with respect to the energy norm shall be estimated. Note that, in contrast to related work on this topic (see, e. g., [We]), we suppose the *solution* u to be a member of F, not the *right-hand side* or *data* f. A link between those two approaches will be provided in Section 6. Furthermore, note that we deal with linear information in the following. As a classical result, it is well known that for f in the standard Sobolev space of the smoothness class r, i. e. $f \in F := H^r(\Omega)$, the ε-complexity of this class of problems is proportional to $\varepsilon^{-d/r}$, making it intractable in the sense of [TWW], since the work that is necessary to

obtain an ε-approximation of the solution grows exponentially in d. This is called the *curse of dimension.*
The growth in complexity is surely one reason why it is a nontrivial task to extend a well developed 2 D program code to the 3 D case. Usually, both storage requirements and run times grow tremendously. Thus, for 3 D applications such as problems from fluid dynamics or structural mechanics, the most elaborate solver routines like multigrid or multilevel methods must be applied in order to be able to obtain solutions of a sufficient accuracy. Additionally, the code must be implemented on a high-performance computer. For problems of dimension $d > 3$, things become even worse. Schrödinger's equation with $p > 1$ particles and, consequently, $d = 3p$ dimensions, or similar applications from physics, still can not be dealt with properly. For such problems, it is a common approach to fall back upon simplified sets of equations of a moderate dimension $d = 3$ and, thus, to settle for an only crude approximation of the interesting quantities such as wave functions and eigenvalues in the Schrödinger equation case. Another possibility for tackling high-dimensional problems is to switch over to pseudorandom or Monte Carlo methods. Originally having been studied for purposes of numerical quadrature, this approach, in the meantime, has become a widespread strategy for high-dimensional PDEs, too, especially on supercomputers. In spite of rather slow convergence, pseudorandom methods have turned out to be very attractive due to their robustness with respect to d: Their complexity is usually proportional to ε^{-2} and, thus, is independent of d. However, convergence results hold only in a probabilistic setting (central limit theorem).
In any case, we have to follow the well-known rule of numerical mathematics: *If you can not solve your problem, change it.* However, since we are not willing to use another mathematical model, but want to stick to our given PDE of specified dimension, the only chance is to look for a solution in a different class of function spaces instead. These considerations have been the starting point of the sparse grid concept [Ze, Gr1, Bu1, Bu2, GSZ, GZZ, BD], where spaces with continuous or (in some sense) bounded mixed derivatives are taken into account. Similar spaces have been recently studied in [Te, WW]. Using a hierarchical subspace decomposition based on tensor product type hierarchical bases of linear splines, and omitting "unimportant" subspaces that possess a relatively large dimension but reduce the error only slightly, it can be shown [Ze, Bu1, Bu2, GSZ] that the number of degrees of freedom and, thus, the number of operations to solve the problem with a multigrid method is proportional to $h^{-1} \cdot |\log h|^{d-1}$, whereas the achieved accuracy is of the order $O(h^2 \cdot |\log h|^{d-1})$ with respect to the L_2- or L_∞-norm and $O(h)$ with respect to the energy norm. Here, h denotes the smallest mesh width occurring in a regular sparse grid. Consequently, for the L_2- and L_∞-norm, we obtain an upper bound on the ε-complexity of the order $O(\varepsilon^{-\frac{1}{2}} \cdot |\log \varepsilon|^{\frac{3}{2}\cdot(d-1)})$, and for the energy norm, we are able to come up with an upper bound on the ε-complexity proportional to $\varepsilon^{-1} \cdot |\log \varepsilon|^{d-1}$.
Note that the sparse grid scheme presented first in [Ze, Gr1] for PDEs has already been known for several years in interpolation, approximation, and recovery theory as well as in numerical quadrature under the different names "hyperbolic crosses" [Bab], "Boolean methods" [Del], and "discrete blending" [Bas, Go, HM]. A general framework of reduced tensor product approximation spaces has been presented by Smolyak [Sm] (cf. also [WW]).
In this paper, we push the ideas underlying the sparse grid approach a little bit further. Considering the cost-accuracy ratio of the different subspaces involved in the hierarchical subspace splitting of tensor product type that leads to the definition of (standard) sparse grids [Ze, Bu1, Bu2], we follow [Bu1] and propose another approximation space that is even more reduced or "sparser" and that involves only $O(h^{-1})$ grid points or degrees of freedom, respectively. It turns out that the inter-

polation error in this subspace of the underlying Sobolev space is of order $O(h)$ with respect to the energy norm. Since, as usual in finite element proofs, the interpolation error is an upper bound for the approximation error in the subspace, and since there is a multilevel preconditioner based on prewavelets [GO2] with a condition number of the resulting system matrix of $O(1)$, we can construct a nested iteration method with a complexity of $O(\varepsilon^{-1})$ for arbitrary dimension d, provided that we restrict ourselves to the Sobolev space of bounded mixed derivatives up to second order in each direction. As a result, we get a practically usable hierarchical basis finite element method with piecewise d-linear ansatz functions for the Laplacian of an ε-complexity of $O(\varepsilon^{-1})$.

The outline of this paper is as follows: In Section 2, we present the hierarchical subspace splitting based on the tensor product approach and, besides some notation, provide the basic estimates for the number of grid points involved in each subspace and for the subspaces' contribution to an interpolant of some function u with respect to various norms. Section 3 discusses two different choices of finite dimensional approximation spaces (i. e., finite dimensional subspaces of the underlying Sobolev space H^1 or H_0^1, respectively) — the classical full grid space and the standard sparse grid space, and gives the order of the corresponding interpolation error. In Section 4, we study the cost-accuracy relation of the different subspaces involved and derive a sparse grid space based on the energy norm. We give an estimate for the number of degrees of freedom involved, and we derive an upper bound for the interpolation error with respect to the energy norm. Section 5 deals with practical aspects of the multilevel solver. The discussion in Section 6 embeds our results into recent related work on this topic. Finally, we give some concluding remarks.

2 Multilevel subspace splitting

Let $\bar{\Omega} := [0,1]^d$ be the d-dimensional unit cube, and let $\mathbf{x} := (x_1, ..., x_d)$. On $\bar{\Omega}$, we consider multivariate functions with (in some sense) bounded weak mixed derivatives

$$D^{\boldsymbol{\alpha}} u \;:=\; \frac{\partial^{|\boldsymbol{\alpha}|_1} u}{\partial x_1^{\alpha_1} ... \partial x_d^{\alpha_d}} \tag{1}$$

up to $|\boldsymbol{\alpha}|_1 = 2d$. Here and in the following, $\boldsymbol{\alpha} \in \mathbb{N}_0^d$ denotes a d-dimensional multi-index. We write $\mathbf{0}$ for $(0, \ldots, 0)$ and $\mathbf{1}$ for $(1, \ldots, 1)$. Furthermore, let $\leq$ etc. denote the corresponding elementwise relations, and let $|\boldsymbol{\alpha}|_1 := \sum_{j=1}^d \alpha_j$ and $|\boldsymbol{\alpha}|_\infty := \max_{1 \leq j \leq d} \alpha_j$ denote the discrete L_∞- and the discrete L_1-norm of $\boldsymbol{\alpha}$, respectively. For $p = \infty$ and $p = 2$, we study the spaces

$$X^p(\Omega) \;:=\; \left\{ u : \mathbb{R}^d \supset \bar{\Omega} \to \mathbb{R},\; D^{\boldsymbol{\alpha}} u \in L_p(\Omega),\; |\boldsymbol{\alpha}|_\infty \leq 2 \right\}, \tag{2}$$

together with their homogeneous counterparts $X_0^p(\Omega)$ of functions vanishing on the boundary, which will be of interest throughout this paper. For functions $u \in X_0^\infty(\Omega)$ or $u \in X_0^2(\Omega)$, respectively, we introduce the semi-norms

$$\begin{aligned} |u|_\infty &:= \left\| D^{\mathbf{2}} u \right\|_\infty , \\ |u|_2 &:= \left\| D^{\mathbf{2}} u \right\|_2 \;=\; \left(\int_\Omega \left| D^{\mathbf{2}} u(\mathbf{x}) \right|^2 \, \mathrm{d}\mathbf{x} \right)^{1/2} . \end{aligned} \tag{3}$$

Now, with the multi-index $\mathbf{l} = (l_1, ..., l_d) \in \mathbb{N}^d$, we consider the family of standard rectangular grids

$$\{\Omega_{\mathbf{l}}, \mathbf{l} \in \mathbb{N}^d\} \tag{4}$$

on $\bar{\Omega}$ with mesh size $h_{\mathbf{l}} := (h_{l_1}, ..., h_{l_d}) := (2^{-l_1}, ..., 2^{-l_d})$. That is, $\Omega_{\mathbf{l}}$ is equidistant with respect to each coordinate direction, but, in general, has different mesh sizes in the different coordinate directions. The grid points contained in a grid $\Omega_{\mathbf{l}}$ are the points

$$x_{\mathbf{l},\mathbf{i}} \ := \ (x_{l_1,i_1}, ..., x_{l_d,i_d}) \tag{5}$$

with $x_{l_j,i_j} := i_j \cdot h_{l_j} = i_j \cdot 2^{-l_j}$, $i_j = 0, ..., 2^{l_j}$. On each grid $\Omega_{\mathbf{l}}$ or, to be precise, on its respective inner grid points, we define the space $V_{\mathbf{l}}$ of piecewise d-linear functions,

$$V_{\mathbf{l}} \ := \ \text{span}\{\phi_{\mathbf{l},\mathbf{i}}, i_j = 1, ..., 2^{l_j} - 1, j = 1, ..., d\}\,, \tag{6}$$

which is spanned by the usual d-dimensional piecewise d-linear hat functions

$$\phi_{\mathbf{l},\mathbf{i}}(\mathbf{x}) \ := \ \prod_{j=1}^{d} \phi_{l_j,i_j}(x_j)\,. \tag{7}$$

Here, the one-dimensional functions $\phi_{l_j,i_j}(x_j)$ with support $[x_{l_j,i_j} - h_{l_j}, x_{l_j,i_j} + h_{l_j}] = [(i_j - 1)h_{l_j}, (i_j + 1)h_{l_j}]$ can be created from a unique mother function $\phi(x)$,

$$\phi(x) \ := \ \begin{cases} 1 - |x| & \text{if } x \ \in \ ]-1, 1[\,, \\ \quad 0 & \text{otherwise,} \end{cases} \tag{8}$$

by dilation and translation, i. e.

$$\phi_{l_j,i_j}(x_j) \ := \ \phi\Big(\frac{x_j - i_j \cdot h_{l_j}}{h_{l_j}}\Big)\,. \tag{9}$$

In the previous definitions and in the following, the multi-index $\mathbf{l} \in \mathbb{N}^d$ indicates the level of a grid or a space or a function, respectively, whereas the multi-index $\mathbf{i} \in \mathbb{N}^d$ denotes the location of a given grid point $x_{\mathbf{l},\mathbf{i}}$ or of the respective basis function $\phi_{\mathbf{l},\mathbf{i}}(\mathbf{x})$.

Now, we can define the difference spaces

$$W_{\mathbf{l}} \ := \ V_{\mathbf{l}} \ominus \sum_{j=1}^{d} V_{\mathbf{l}-\mathbf{e}_j}\,, \tag{10}$$

where $\mathbf{e}_j$ denotes the j-th unit vector. To complete this definition, we formally set $V_{\mathbf{l}} := \emptyset$ if $l_j = 0$ for at least one $j \in \{1, ..., d\}$. These hierarchical difference spaces allow us the definition of a multilevel subspace splitting, i. e. the definition of the space V as a direct sum of subspaces,

$$V \ := \ \sum_{l_1=1}^{\infty} \cdots \sum_{l_d=1}^{\infty} W_{(l_1,..l_d)} \ = \ \bigoplus_{\mathbf{l}\in\mathbb{N}^d} W_{\mathbf{l}}\,, \tag{11}$$

which is, except for completion with respect to the H^1-norm, the underlying Sobolev space, i. e. $\bar{V} = H_0^1(\Omega)$. In the next section, we will deal with finite dimensional subspaces of $H_0^1(\Omega)$. Note that, e. g., the choice

$$V_n^{(\infty)} \ := \ \bigoplus_{|\mathbf{l}|_\infty \le n} W_{\mathbf{l}}\,, \tag{12}$$

provides a sequence of nested finite dimensional approximation spaces $V_n^{(\infty)} \subset V_{n+1}^{(\infty)}$ with

$$\bigcup_{n=1}^{\infty} V_n^{(\infty)} = V\,. \tag{13}$$

As it can be easily seen from (6) and (10), the introduction of index sets $\mathbf{I_l}$,

$$\mathbf{I_l} \;:=\; \{(i_1, ..., i_d) \in \mathbb{N}^d, i_j = 1, ..., 2^{l_j} - 1, i_j \text{ odd}, j = 1, ..., d\}\,, \tag{14}$$

leads to

$$W_\mathbf{l} \;=\; \operatorname{span}\{\phi_{\mathbf{l},\mathbf{i}}, \mathbf{i} \in \mathbf{I_l}\}\,. \tag{15}$$

Therefore, the family of functions

$$\{\phi_{\mathbf{l},\mathbf{i}}, \mathbf{i} \in \mathbf{I_l}, \mathbf{l} \in \mathbb{N}^d\} \tag{16}$$

is just a hierarchical basis [Fa, Ys1, Ys2] of $H_0^1(\Omega)$ that generalizes the 1 D hierarchical basis of [Fa] to the d-dimensional case by means of a tensor product approach. Note that the supports of all basis functions $\phi_{\mathbf{l},\mathbf{i}}(\mathbf{x})$ in (15) spanning $W_\mathbf{l}$ are mutually disjoint. Now, any function $u \in H_0^1(\Omega)$ can be split accordingly by

$$u(\mathbf{x}) \;=\; \sum_\mathbf{l} u_\mathbf{l}(\mathbf{x}), \quad u_\mathbf{l} \in W_\mathbf{l}, \tag{17}$$

and

$$u_\mathbf{l}(\mathbf{x}) \;=\; \sum_{\mathbf{i} \in \mathbf{I_l}} v_{\mathbf{l},\mathbf{i}} \cdot \phi_{\mathbf{l},\mathbf{i}}(\mathbf{x}), \tag{18}$$

where $v_{\mathbf{l},\mathbf{i}} \in \mathbb{R}$ are the coefficient values of the hierarchical product basis representation of u. Obviously, since (16) forms a basis of V or $\bar{V}$, respectively, all hierarchical coefficients $v_{\mathbf{l},\mathbf{i}}$ are determined uniquely.
After these preliminaries, we are now able to state the following five lemmata concerning the number of unknowns involved in each subspace, the values of different norms of the basis functions $\phi_{\mathbf{l},\mathbf{i}}(\mathbf{x})$, the characteristics and size of the hierarchical coefficients $v_{\mathbf{l},\mathbf{i}}$, and the norms of the contributions $u_\mathbf{l}$ of $W_\mathbf{l}$ according to (18).

Lemma 1 *The dimension of the subspace $W_\mathbf{l}$ is*

$$|W_\mathbf{l}| = 2^{|\mathbf{l}-\mathbf{1}|_1}. \tag{19}$$

*P*roof: (19) follows immediately from (14) and (15). □

Following the finite element terminology, let the energy norm of $u \in X_0^p(\Omega)$, $p \in \{2, \infty\}$, be defined as

$$||u||_E := \Big(\int_\Omega \sum_{j=1}^d \Big(\frac{\partial u(x)}{\partial x_j} \Big)^2 \mathrm{d}\mathbf{x} \Big)^{1/2}. \tag{20}$$

Lemma 2 *For any basis function $\phi_{\mathbf{l},\mathbf{i}}(\mathbf{x})$, the following equations hold:*

$$\begin{aligned} ||\phi_{\mathbf{l},\mathbf{i}}||_\infty \;&=\; 1\,, \\ ||\phi_{\mathbf{l},\mathbf{i}}||_p \;&=\; \Big(\frac{2}{p+1}\Big)^{d/p} \cdot 2^{-|\mathbf{l}|_1/p} \qquad (p \geq 1)\,, \\ ||\phi_{\mathbf{l},\mathbf{i}}||_E \;&=\; \sqrt{2} \cdot \Big(\frac{2}{3}\Big)^{(d-1)/2} \cdot 2^{-|\mathbf{l}|_1/2} \cdot \Big(\sum_{j=1}^d 2^{2l_j}\Big)^{1/2}. \end{aligned} \tag{21}$$

*P*roof: Straightforward calculation (cf. [Bu1, Bu2]). □

Now, we consider the hierarchical coefficient values $v_{\mathbf{l},\mathbf{i}}$ in more detail. They can be computed from the function values $u(\mathbf{x}_{\mathbf{l},\mathbf{i}})$ in the following way:

$$v_{\mathbf{l},\mathbf{i}} \;=\; \Big(\prod_{j=1}^{d} \big[\; -\tfrac{1}{2} \quad 1 \quad -\tfrac{1}{2} \;\big]_{x_{l_j,i_j},l_j}\Big) u \;=:\; \Big(\prod_{j=1}^{d} I_{x_{l_j,i_j},l_j}\Big) u \;=:\; I_{\mathbf{x}_{\mathbf{l},\mathbf{i}},\mathbf{l}} u \,. \qquad (22)$$

This follows from the definition of the spaces $W_{\mathbf{l}}$ and their basis functions (15). Here, as usual in multigrid terminology, $I_{\mathbf{x}_{\mathbf{l},\mathbf{i}},\mathbf{l}}$ denotes a d-dimensional stencil which gives the coefficients for a linear combination of nodal values of u, see also [Ha1].

Lemma 3 *Let $\psi_{l_j,i_j}(x_j) := -2^{-(l_j+1)} \cdot \phi_{l_j,i_j}(x_j)$, and let $\psi_{\mathbf{l},\mathbf{i}}(\mathbf{x}) := \prod_{j=1}^{d} \psi_{l_j,i_j}(x_j)$. For any coefficient value $v_{\mathbf{l},\mathbf{i}}$ of the hierarchical representation (17) and (18) of $u \in X_0^p(\Omega)$, $p \in \{2,\infty\}$, the following relation holds:*

$$v_{\mathbf{l},\mathbf{i}} \;=\; \int_{\Omega} \psi_{\mathbf{l},\mathbf{i}}(\mathbf{x}) \cdot D^{\mathbf{2}} u(\mathbf{x}) \; \mathrm{d}\mathbf{x} \,. \qquad (23)$$

*P*roof: First, we look at the simplest case $d = 1$. Partial integration provides

$$\begin{aligned}
&\int_{\Omega} \psi_{l_j,i_j}(x_j) \cdot \frac{\partial^2 u(x_j)}{\partial x_j^2} \; \mathrm{d}\mathbf{x} \\
&= \int_{x_{l_j,i_j}-h_{l_j}}^{x_{l_j,i_j}+h_{l_j}} \psi_{l_j,i_j}(x_j) \cdot \frac{\partial^2 u(x_j)}{\partial x_j^2} \; \mathrm{d}x_j \\
&= \left[\psi_{l_j,i_j}(x_j) \cdot \frac{\partial u(x_j)}{\partial x_j}\right]_{x_{l_j,i_j}-h_{l_j}}^{x_{l_j,i_j}+h_{l_j}} \;-\; \int_{x_{l_j,i_j}-h_{l_j}}^{x_{l_j,i_j}+h_{l_j}} \frac{\partial \psi_{l_j,i_j}(x_j)}{\partial x_j} \cdot \frac{\partial u(x_j)}{\partial x_j} \; \mathrm{d}x_j \\
&= \int_{x_{l_j,i_j}-h_{l_j}}^{x_{l_j,i_j}} \frac{1}{2} \cdot \frac{\partial u(x_j)}{\partial x_j} \; \mathrm{d}x_j \;-\; \int_{x_{l_j,i_j}}^{x_{l_j,i_j}+h_{l_j}} \frac{1}{2} \cdot \frac{\partial u(x_j)}{\partial x_j} \; \mathrm{d}x_j \\
&= I_{x_{l_j,i_j},l_j} u \,,
\end{aligned}$$

since $\psi_{l_j,i_j}(x_{l_j,i_j} - h_{l_j}) = \psi_{l_j,i_j}(x_{l_j,i_j} + h_{l_j}) = 0$ and since $\partial \psi_{l_j,i_j}(x_j)/\partial x_j \in \{\frac{1}{2}, -\frac{1}{2}\}$ due to the construction of ψ_{l_j,i_j} and ϕ_{l_j,i_j}. The tensor product approach leads to a straightforward generalization to $d > 1$, which shows (23) in a strong sense. For our case of weak derivatives, we get

$$v_{\mathbf{l},\mathbf{i}} \;=\; \int_{\Omega} \psi_{\mathbf{l},\mathbf{i}}(\mathbf{x}) \cdot D^{\mathbf{2}} u(\mathbf{x}) \; \mathrm{d}\mathbf{x} \;=\; \int_{\Omega} \hat{\psi}_{\mathbf{l},\mathbf{i}}(\mathbf{x}) \cdot u(\mathbf{x}) \; \mathrm{d}\mathbf{x} \,,$$

where $\hat{\psi}_{\mathbf{l},\mathbf{i}}(\mathbf{x})$ equals $\psi_{\mathbf{l},\mathbf{i}}(\mathbf{x})$ in the weak sense and is just a linear combination of 3^d Dirac pulses of the weights $(-1/2, 1, -1/2)$ in each direction. □

Lemma 4 *Let $u \in X_0^p(\Omega)$, $p \in \{2,\infty\}$, be given in the representation (17) and (18). Then, the following estimates for the hierarchical coefficient values $v_{\mathbf{l},\mathbf{i}}$ hold:*

$$\begin{aligned}
|v_{\mathbf{l},\mathbf{i}}| \;&\leq\; 2^{-d} \cdot 2^{-2 \cdot |\mathbf{l}|_1} \cdot |u|_{\infty} \,, \\
|v_{\mathbf{l},\mathbf{i}}| \;&\leq\; 2^{-d} \cdot \left(\frac{2}{3}\right)^{d/2} \cdot 2^{-3/2 \cdot |\mathbf{l}|_1} \cdot \left\|\chi_{\mathbf{l},\mathbf{i}} \cdot D^{\mathbf{2}} u\right\|_2 \,,
\end{aligned} \qquad (24)$$

where $\chi_{\mathbf{l},\mathbf{i}}(\mathbf{x})$ is the characteristic function of the support of $\phi_{\mathbf{l},\mathbf{i}}(\mathbf{x})$.

Proof: With (23), (21), and with the definition of $\psi_{\mathbf{l},\mathbf{i}}$, we get

$$\begin{aligned}|v_{\mathbf{l},\mathbf{i}}| \;=\; \left|\int_\Omega \psi_{\mathbf{l},\mathbf{i}}(\mathbf{x})\cdot D^{\mathbf{2}}u(\mathbf{x})\,\mathrm{d}\mathbf{x}\right| \;&\leq\; \|\psi_{\mathbf{l},\mathbf{i}}\|_1 \cdot \left\|\chi_{\mathbf{l},\mathbf{i}}\cdot D^{\mathbf{2}}u\right\|_\infty \\ &=\; 2^{-d}\cdot 2^{-|\mathbf{l}|_1}\cdot \|\phi_{\mathbf{l},\mathbf{i}}\|_1 \cdot |\chi_{\mathbf{l},\mathbf{i}}\cdot u|_\infty \\ &\leq\; 2^{-d}\cdot 2^{-2\cdot|\mathbf{l}|_1}\cdot |u|_\infty\,.\end{aligned}$$

For the second bound, the Cauchy-Schwarz inequality provides

$$\begin{aligned}|v_{\mathbf{l},\mathbf{i}}| \;=\; \left|\int_\Omega \psi_{\mathbf{l},\mathbf{i}}(\mathbf{x})\cdot D^{\mathbf{2}}u(\mathbf{x})\,\mathrm{d}\mathbf{x}\right| \;&\leq\; \|\psi_{\mathbf{l},\mathbf{i}}\|_2 \cdot \left\|\chi_{\mathbf{l},\mathbf{i}}\cdot D^{\mathbf{2}}u\right\|_2 \\ &=\; 2^{-d}\cdot 2^{-|\mathbf{l}|_1}\cdot \|\phi_{\mathbf{l},\mathbf{i}}\|_2 \cdot \left\|\chi_{\mathbf{l},\mathbf{i}}\cdot D^{\mathbf{2}}u\right\|_2\,,\end{aligned}$$

which, with (21), is the desired result. □

Lemma 5 *Let $u \in X_0^p(\Omega)$, $p \in \{2,\infty\}$, be given in the representation (17) and (18). Then, the following estimates for its components $u_{\mathbf{l}} \in W_{\mathbf{l}}$ hold:*

$$\begin{aligned}||u_{\mathbf{l}}||_\infty \;&\leq\; 2^{-d}\cdot 2^{-2\cdot|\mathbf{l}|_1}\cdot |u|_\infty\,, \\ ||u_{\mathbf{l}}||_2 \;&\leq\; 3^{-d}\cdot 2^{-2\cdot|\mathbf{l}|_1}\cdot |u|_2\,, \\ ||u_{\mathbf{l}}||_E \;&\leq\; \frac{1}{2\cdot 12^{(d-1)/2}}\cdot 2^{-2\cdot|\mathbf{l}|_1}\cdot\left(\sum_{j=1}^d 2^{2\cdot l_j}\right)^{1/2}\cdot |u|_\infty\,, \\ ||u_{\mathbf{l}}||_E \;&\leq\; \sqrt{3}\cdot 3^{-d}\cdot 2^{-2\cdot|\mathbf{l}|_1}\cdot\left(\sum_{j=1}^d 2^{2\cdot l_j}\right)^{1/2}\cdot |u|_2\,.\end{aligned} \tag{25}$$

Proof: Since the supports of all basis functions $\phi_{\mathbf{l},\mathbf{i}}$ contributing to $u_{\mathbf{l}}$ in the representation (18) are mutually disjoint, the first estimate follows immediately from the respective statements in (21) and (24). For the estimate concerning the L_2-norm, we get with the same argument of disjoint supports and with (21) and (24)

$$\begin{aligned}||u_{\mathbf{l}}||_2^2 \;=\; \left\|\sum_{\mathbf{i}\in\mathbf{I}_{\mathbf{l}}} v_{\mathbf{l},\mathbf{i}}\cdot\phi_{\mathbf{l},\mathbf{i}}\right\|_2^2 \;&=\; \sum_{\mathbf{i}\in\mathbf{I}_{\mathbf{l}}} |v_{\mathbf{l},\mathbf{i}}|^2\cdot||\phi_{\mathbf{l},\mathbf{i}}||_2^2 \\ &\leq\; \sum_{\mathbf{i}\in\mathbf{I}_{\mathbf{l}}} \frac{1}{6^d}\cdot 2^{-3\cdot|\mathbf{l}|_1}\cdot\left\|\chi_{\mathbf{l},\mathbf{i}}\cdot D^{\mathbf{2}}u\right\|_2^2\cdot\left(\frac{2}{3}\right)^d\cdot 2^{-|\mathbf{l}|_1} \\ &=\; 9^{-d}\cdot 2^{-4\cdot|\mathbf{l}|_1}\cdot\left\|D^{\mathbf{2}}u\right\|_2^2\,.\end{aligned}$$

Finally, analogous calculations based on the disjoint supports and on the corresponding relations from (21) and (24) lead to the respective estimates for $||u_{\mathbf{l}}||_E$. □

Concerning the energy estimates in Lemma 5, note that the order of $||u_{\mathbf{l}}||_E$ may depend on which semi-norm is used for the bounds of $|v_{\mathbf{l},\mathbf{i}}|$ (cf. Lemma 4). For both $|.|_\infty$ and $|.|_2$, we get $2^{-2\cdot|\mathbf{l}|_1}$. However, note that for $|.|_1$ defined in an analogous way, only a bound of the order $2^{-3/2\cdot|\mathbf{l}|_1}$ can be given.

3 Interpolation in finite dimensional spaces

We now turn to finite dimensional subspaces of $H_0^1(\Omega)$. With the help of the hierarchical multilevel splitting (11) introduced in the previous section, different subspaces can be characterized easily using the level multi-index $\mathbf{l}$. To this end, we basically switch from taking into account all levels $\mathbf{l} \geq \mathbf{1}$ to finite sets of active levels in the summation (17). For a given parameter $n \in \mathbb{N}$, e. g., one possible choice has already been introduced in (12):

$$V_n^{(\infty)} = \bigoplus_{|\mathbf{l}|_\infty \leq n} W_{\mathbf{l}} . \tag{26}$$

The finite dimensional subspace $V_n^{(\infty)}$ is just the usual space of piecewise d-linear functions on the regular grid $\Omega_{n\cdot\mathbf{1}}$ with equidistant mesh size $h_n = 2^{-n}$ in each coordinate direction. It is a well-known fact that the dimension of this space (i. e., the number of inner grid points in the underlying grid — note that we work with subspaces of $H_0^1(\Omega)$) is

$$\left| V_n^{(\infty)} \right| = (2^n - 1)^d = O(2^{d\cdot n}) . \tag{27}$$

We have the following results for the error $u - u_n^{(\infty)}$ of the interpolant $u_n^{(\infty)}$ of u in $V_n^{(\infty)}$:

Lemma 6 *For $u \in X_0^p(\Omega)$, $p \in \{2, \infty\}$, the following estimates for the different norms of $u - u_n^{(\infty)}$ hold:*

$$\begin{aligned}
||u - u_n^{(\infty)}||_\infty &\leq \frac{d}{6^d} \cdot 2^{-2n} \cdot |u|_\infty \;=\; O(h_n^2) , \\
||u - u_n^{(\infty)}||_2 &\leq \frac{d}{9^d} \cdot 2^{-2n} \cdot |u|_2 \;=\; O(h_n^2) , \\
||u - u_n^{(\infty)}||_E &\leq \frac{d^{3/2}}{2 \cdot 3^{(d-1)/2} \cdot 6^{d-1}} \cdot 2^{-n} \cdot |u|_\infty \;=\; O(h_n) , \\
||u - u_n^{(\infty)}||_E &\leq \frac{d^{3/2}}{\sqrt{3} \cdot 9^{d-1}} \cdot 2^{-n} \cdot |u|_2 \;=\; O(h_n) .
\end{aligned} \tag{28}$$

Proof: We show the bound for the L_∞-norm. From (25), we get

$$\begin{aligned}
||u - u_n^{(\infty)}||_\infty \leq \sum_{|\mathbf{l}|_\infty > n} ||u_{\mathbf{l}}||_\infty &\leq \frac{1}{2^d} \cdot |u|_\infty \cdot \sum_{|\mathbf{l}|_\infty > n} 2^{-2\cdot|\mathbf{l}|_1} \\
&= \frac{1}{2^d} \cdot |u|_\infty \cdot \Big(\sum_{\mathbf{l}} 4^{-|\mathbf{l}|_1} - \sum_{|\mathbf{l}|_\infty \leq n} 4^{-|\mathbf{l}|_1} \Big) \\
&= \frac{1}{2^d} \cdot |u|_\infty \cdot \Big(\frac{1}{3^d} - \big(\sum_{i=1}^{n} 4^{-i} \big)^d \Big) \\
&= \frac{1}{2^d} \cdot |u|_\infty \cdot \frac{1}{3^d} \cdot \big(1 - (1 - 4^{-n})^d \big) \\
&\leq \frac{d}{6^d} \cdot |u|_\infty \cdot 4^{-n} .
\end{aligned}$$

The L_2-proof can be obtained in the same way. For the error with respect to the energy norm, (25) leads to

$$\begin{aligned}
||u - u_n^{(\infty)}||_E &\leq \frac{d^{1/2}}{2 \cdot 12^{(d-1)/2}} \cdot |u|_\infty \cdot \sum_{|\mathbf{l}|_\infty > n} 2^{-2\cdot|\mathbf{l}|_1} \cdot \max_{1 \leq j \leq d} 2^{l_j} \\
&\leq \frac{d^{3/2}}{2 \cdot 12^{(d-1)/2}} \cdot |u|_\infty \cdot \sum_{|\mathbf{l}|_\infty = l_1 > n} 2^{-2\cdot|\mathbf{l}|_1} \cdot 2^{l_1}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{d^{3/2}}{2\cdot 12^{(d-1)/2}}\cdot |u|_\infty \cdot \sum_{l_1>n} 2^{-l_1}\cdot \Big(\sum_{l_j=1}^{l_1} 4^{-l_j}\Big)^{d-1} \\
&\le \frac{d^{3/2}}{2\cdot 12^{(d-1)/2}}\cdot |u|_\infty \cdot \frac{1}{3^{d-1}}\cdot 2^{-n}\,,
\end{aligned}
$$

and an analogous argument provides the second bound. □

Note that our regularity assumptions differ from those of standard finite element theory (cf. [SF], e. g.), but without effects on the order of accuracy.
If we look at the cost and at the benefit of a certain hierarchical subspace, two observations are crucial. First, (19) indicates that subspaces $W_{\mathbf{l}}$ with fixed $|\mathbf{l}|_1$ cost the same, i. e., they involve the same number of grid points or degrees of freedom. Second, we learn from (25) that the benefit of all $W_{\mathbf{l}}$ with fixed $|\mathbf{l}|_1$, i. e. their contribution to the interpolant with respect to the L_2- or the L_∞-norm, is of the same order of magnitude, too. Therefore, it turns out to be more reasonable to define finite dimensional subspaces of $H_0^1(\Omega)$ based on $|\mathbf{l}|_1$ rather than on $|\mathbf{l}|_\infty$ as in (12):

$$
V_n^{(1)} \;:=\; \bigoplus_{|\mathbf{l}|_1\le n+d-1} W_{\mathbf{l}}. \tag{29}
$$

The grids corresponding to the approximation spaces $V_n^{(1)}$ are called sparse grids and have been studied in detail in [Ze, Bu1, Bu2, Gr1, Gr2, GSZ, GZZ, BD, BDZ]. Examples of sparse grids (a regular two-dimensional and an adaptively refined three-dimensional one) are given in Figure 1.

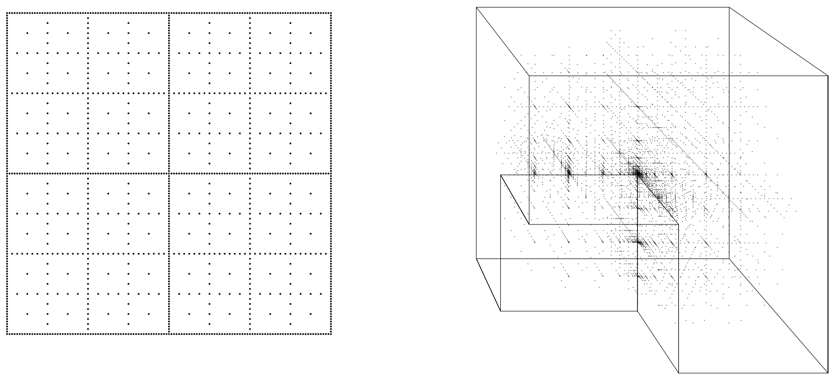

Figure 1: Sparse grids: regular example (left) and adaptive one (right)

The following lemma provides a formula for the number of inner grid points in a sparse grid.

Lemma 7 *The dimension of the space $V_n^{(1)}$ is given by*

$$
\begin{aligned}
\left|V_n^{(1)}\right| &= \sum_{i=0}^{n-1} 2^i\cdot\binom{d-1+i}{d-1} \\
&= (-1)^d \;+\; 2^n\cdot\sum_{i=0}^{d-1}\binom{n+d-1}{i}\cdot(-2)^{d-1-i} \qquad (30)\\
&= 2^n\cdot\Big(\frac{n^{d-1}}{(d-1)!}+O(n^{d-2})\Big).
\end{aligned}
$$

*P*roof: With (19) and (29), we get

$$\left|V_n^{(1)}\right| = \Big|\bigoplus_{|\mathbf{l}|_1 \leq n+d-1} W_{\mathbf{l}}\Big| = \sum_{|\mathbf{l}|_1 \leq n+d-1} 2^{|\mathbf{l}-\mathbf{1}|_1} = \sum_{i=d}^{n+d-1} 2^{i-d} \sum_{|\mathbf{l}|_1 = i} 1$$
$$= \sum_{i=d}^{n+d-1} 2^{i-d} \cdot \binom{i-1}{d-1}$$
$$= \sum_{i=0}^{n-1} 2^i \cdot \binom{d-1+i}{d-1},$$

since there are $\binom{i-1}{d-1}$ possibilities to form the sum i with d natural numbers. Furthermore,

$$\sum_{i=0}^{n-1} 2^i \cdot \binom{d-1+i}{d-1}$$
$$= \frac{1}{(d-1)!} \cdot \sum_{i=0}^{n-1} \left(x^{i+d-1}\right)^{(d-1)} \Big|_{x=2}$$
$$= \frac{1}{(d-1)!} \cdot \left(x^{d-1} \cdot \frac{1-x^n}{1-x}\right)^{(d-1)} \Big|_{x=2}$$
$$= \frac{1}{(d-1)!} \cdot \sum_{i=0}^{d-1} \binom{d-1}{i} \cdot \left(x^{d-1} - x^{n+d-1}\right)^{(i)} \cdot \left(\frac{1}{1-x}\right)^{(d-1-i)} \Big|_{x=2}$$
$$= (-1)^d + 2^n \cdot \sum_{i=0}^{d-1} \binom{n+d-1}{i} \cdot (-2)^{d-1-i},$$

from which the statement concerning the order and the leading coefficient follows immediately. □

For a comparison of the cost of full grid spaces $V_n^{(\infty)}$ and sparse grid spaces $V_n^{(1)}$, (30) has to be compared with (27). On the other hand, we have to look at the interpolation accuracy that can be obtained on sparse grids. For that, we deal with the interpolation error $u - u_n^{(1)}$ of the sparse grid interpolant $u_n^{(1)} \in V_n^{(1)}$ which, due to (17) and (29), can be written as

$$u - u_n^{(1)} = \sum_{\mathbf{l}} u_{\mathbf{l}} - \sum_{|\mathbf{l}|_1 \leq n+d-1} u_{\mathbf{l}} = \sum_{|\mathbf{l}|_1 > n+d-1} u_{\mathbf{l}}.$$

Therefore, for any norm $||.||$, we have

$$||u - u_n^{(1)}|| \leq \sum_{|\mathbf{l}|_1 > n+d-1} ||u_{\mathbf{l}}||. \tag{31}$$

The following Lemma 8 will conclude the preliminaries for the estimates of the interpolation error with respect to the different norms we are interested in. First, we define for $d, n \in \mathbb{N}$

$$A(d,n) := \sum_{k=0}^{d-1} \binom{n+d-1}{k} = \frac{n^{d-1}}{(d-1)!} + O(n^{d-2}). \tag{32}$$

Lemma 8 *For purposes of summation over all grid points $x_{\mathbf{l},\mathbf{i}}$ with corresponding basis functions $\phi_{\mathbf{l},\mathbf{i}} \notin V_n^{(1)}$, we obtain for arbitrary $s \in \mathbb{N}$*

$$\begin{aligned} \sum_{|\mathbf{l}|_1 > n+d-1} 2^{-s\cdot|\mathbf{l}|_1} &= 2^{-s\cdot n}\cdot 2^{-s\cdot d}\cdot \sum_{i=0}^{\infty} 2^{-s\cdot i}\cdot \binom{n+i+d-1}{d-1} \\ &\leq 2^{-s\cdot n}\cdot 2^{-s\cdot d}\cdot 2\cdot A(d,n)\,. \end{aligned} \tag{33}$$

*P*roof: Similarly to the proof of Lemma 7, we get

$$\begin{aligned} \sum_{|\mathbf{l}|_1 > n+d-1} 2^{-s\cdot|\mathbf{l}|_1} &= \sum_{i=n+d}^{\infty} 2^{-s\cdot i}\cdot \sum_{|\mathbf{l}|_1 = i} 1 \\ &= \sum_{i=n+d}^{\infty} 2^{-s\cdot i}\cdot \binom{i-1}{d-1} \\ &= 2^{-s\cdot n}\cdot 2^{-s\cdot d}\cdot \sum_{i=0}^{\infty} 2^{-s\cdot i}\cdot \binom{n+i+d-1}{d-1}. \end{aligned}$$

Since

$$\begin{aligned} \sum_{i=0}^{\infty} x^i\cdot \binom{n+i+d-1}{d-1} &= \frac{x^{-n}}{(d-1)!}\cdot \Big(\sum_{i=0}^{\infty} x^{n+i+d-1}\Big)^{(d-1)} \\ &= \frac{x^{-n}}{(d-1)!}\cdot \Big(x^{n+d-1}\cdot \frac{1}{1-x}\Big)^{(d-1)} \\ &= \frac{x^{-n}}{(d-1)!}\cdot \sum_{k=0}^{d-1} \binom{d-1}{k}\cdot \Big(x^{n+d-1}\Big)^{(k)}\cdot \Big(\frac{1}{1-x}\Big)^{(d-1-k)} \\ &= \sum_{k=0}^{d-1} \binom{n+d-1}{k}\cdot \Big(\frac{x}{1-x}\Big)^{d-1-k}\cdot \frac{1}{1-x}\,, \end{aligned}$$

we get with $x := 2^{-s}$:

$$\sum_{i=0}^{\infty} 2^{-s\cdot i}\cdot \binom{n+i+d-1}{d-1} \;\leq\; 2\cdot \sum_{k=0}^{d-1} \binom{n+d-1}{k} \;=\; 2\cdot A(d,n)\,.$$

□

Finally, (25) from Lemma 5 and (33) from Lemma 8 lead to the main result of this section.

Theorem 1 *For the L_∞-, the L_2-, and the energy norm, the following upper bounds of the interpolation error of a function $u \in X_0^p(\Omega)$, $p \in \{2,\infty\}$, in the sparse grid space $V_n^{(1)}$ hold:*

$$\begin{aligned} \|u-u_n^{(1)}\|_\infty &\leq \frac{2\cdot|u|_\infty}{8^d}\cdot 2^{-2n}\cdot A(d,n) \;=\; O(h_n^2\cdot n^{d-1})\,, \\ \|u-u_n^{(1)}\|_2 &\leq \frac{2\cdot|u|_2}{12^d}\cdot 2^{-2n}\cdot A(d,n) \;=\; O(h_n^2\cdot n^{d-1})\,, \\ \|u-u_n^{(1)}\|_E &\leq \frac{d\cdot|u|_\infty}{2\cdot 3^{(d-1)/2}\cdot 4^{d-1}}\cdot 2^{-n} \;=\; O(h_n)\,, \\ \|u-u_n^{(1)}\|_E &\leq \frac{d\cdot|u|_2}{\sqrt{3}\cdot 6^{d-1}}\cdot 2^{-n} \;=\; O(h_n)\,. \end{aligned} \tag{34}$$

*P*roof: With (25), (31), and (33) for $s = 2$, we get

$$\begin{aligned} \|u - u_n^{(1)}\|_\infty \;\le\; \sum_{|\mathbf{l}|_1 > n+d-1} \|u_\mathbf{l}\|_\infty \;&\le\; \frac{|u|_\infty}{2^d} \cdot \sum_{|\mathbf{l}|_1 > n+d-1} 2^{-2\cdot|\mathbf{l}|_1} \\ &\le\; \frac{2 \cdot |u|_\infty}{8^d} \cdot 2^{-2n} \cdot A(d,n) \end{aligned}$$

and, analogously, the corresponding result for the L_2-norm. Concerning the energy norm, we have

$$\begin{aligned} \|u - u_n^{(1)}\|_E \;\le\; \sum_{|\mathbf{l}|_1 > n+d-1} \|u_\mathbf{l}\|_E \;&\le\; \frac{|u|_\infty}{2 \cdot 12^{(d-1)/2}} \cdot \sum_{|\mathbf{l}|_1 > n+d-1} 4^{-|\mathbf{l}|_1} \cdot \left(\sum_{j=1}^{d} 4^{l_j} \right)^{1/2} \\ &=\; \frac{|u|_\infty}{2 \cdot 12^{(d-1)/2}} \cdot \sum_{i=n+d}^{\infty} 4^{-i} \cdot \sum_{|\mathbf{l}|_1 = i} \left(\sum_{j=1}^{d} 4^{l_j} \right)^{1/2} \\ &\le\; \frac{|u|_\infty}{2 \cdot 12^{(d-1)/2}} \cdot \sum_{i=n+d}^{\infty} d \cdot 2^{-i} \\ &=\; \frac{d \cdot |u|_\infty}{2 \cdot 3^{(d-1)/2} \cdot 4^{d-1}} \cdot 2^{-n}\,, \end{aligned}$$

because

$$\sum_{|\mathbf{l}|_1 = i} \left(\sum_{j=1}^{d} 4^{l_j} \right)^{1/2} \;\le\; d \cdot 2^i$$

(complete induction with respect to d). The second energy estimate can be obtained in exactly the same way, starting from the respective bound in Lemma 6. □

In recent years, the sparse grid spaces $V_n^{(1)}$ have been studied in detail, the focus being on both theoretical and algorithmic aspects as well as on applications in computational fluid dynamics. For further information, see [BD, BDZ, Gr2, GHZ, GO1, GSZ, GZZ] and the references therein.

Note that, in the definition of the discrete spaces $V_n^{(\infty)}$ and $V_n^{(1)}$ introduced up to now, we used the discrete $|.|_\infty$- and $|.|_1$-norm of the multi-index $\mathbf{l}$ in the summation. Of course, the range of possibilities how to choose subsets of $\{\mathbf{l} \in \mathbb{N}^d\}$ is much larger, resulting in more complicated summation domains and, thus, different grid structures. Such more general choices of finite index subsets will be the focus of the next section.

4 Sparse grids based on the energy norm

In the previous section, sparse grids and their underlying approximation spaces $V_n^{(1)}$ have been introduced according to simple cost-accuracy considerations: Subspaces $W_\mathbf{l}$ with a number of degrees of freedom that is large in relation to the improvement of accuracy induced by them have been removed from $V_n^{(\infty)}$. Now, we want to give a formal motivation for definition (29) and, furthermore, derive more general sparse grid spaces depending on the norm in which the error is measured. Actually, the choice of some kind of an optimal grid is a restricted minimization problem: Find a grid (i.e. an approximation space V_n consisting of certain subspaces $W_\mathbf{l}$) such that the interpolation error or a suitable error bound with respect to the respective norm is minimal for some fixed cost (number of grid points). We can characterize such a V_n via the index set $I := \{\mathbf{l} : W_{\mathbf{l}} \in V_n\}$, and we can restrict the range of possible subspaces $W_{\mathbf{l}}$ to $I_{\max} := \{\mathbf{l} : |\mathbf{l}|_\infty \leq N\}$ with N sufficiently large without loss of generality. If we arrange the indices $\mathbf{l}$ from $I_{\max}$ in some linear order, our optimization problem can be formulated as a classical binary knapsack problem:

$$\max_x c^T x \quad \text{with} \quad a^T x \leq b\,. \tag{35}$$

Here, a and c are both vectors of natural numbers of length $M := (2^N - 1)^d$. The components a_i denote the subspaces' dimension $|W_{\mathbf{l}}|$ (i. e., the cost induced by $W_{\mathbf{l}}$), the c_i are just the squared and suitably scaled upper error bounds of the subspaces' contributions to the overall interpolant according to Lemma 5, and b is nothing but the given bound for the overall work count. Note that, without any influence on the optimization process, we use scaled squares of the upper bounds in order to ensure $c_i \in \mathbb{N}$. The binary vector $x \in \{0,1\}^M$, finally, indicates which subspaces $W_{\mathbf{l}}$ do belong to our approximation space V_n and which ones do not. Now, it is a well-known result from combinatorial optimization that, for *rational* $x_i \in [0,1]$, the following algorithm provides an optimal solution x (which has, in general, nothing to do with the optimal *binary* solution, however; see [MT], e. g.):

(1) rearrange the components such that $c_1/a_1 \geq c_2/a_2 \geq \ldots \geq c_M/a_M$;

(2) define $r := \max\{j : \sum_{i=1}^{j} a_i \leq b\}$;

(3) set $x_1 := \ldots := x_r := 1$; $x_{r+2} := \ldots := x_M := 0$; $x_{r+1} := \left(b - \sum_{i=1}^{r} a_i\right)/a_{r+1}$;

Since we can choose our work count b arbitrarily, $x_{r+1} = 1$ and, thus, $x \in \{0,1\}^M$ can be guaranteed. Consequently, we are able to get an optimal *binary* solution for our knapsack problem (35). Concerning our search for optimal grids, the crucial point is the fact that the overall optimization process on $I_{\max}$ can be reduced to the discussion of the quotients

$$\gamma_{\mathbf{l},||.||} \;:=\; \frac{\text{value}_{||.||}(W_{\mathbf{l}})}{\text{cost}(W_{\mathbf{l}})} \;:=\; \frac{||u_{\mathbf{l}}||^2}{|W_{\mathbf{l}}|} \tag{36}$$

that indicate the cost-benefit or cost-accuracy ratio of a subspace $W_{\mathbf{l}}$ with respect to a given norm $||.||$. That is, for an optimal grid in this sense, it is the appropriate choice to take into account those $W_{\mathbf{l}}$ with the biggest values $\gamma_{\mathbf{l},||.||}$ or those $W_{\mathbf{l}}$, respectively, with $\gamma_{\mathbf{l},||.||}$ bigger than some prescribed threshold $\sigma_{||.||}(n)$, where $\sigma_{||.||}(n)$ is chosen to be of the order of $\gamma_{\bar{\mathbf{l}},||.||}$ with $\bar{\mathbf{l}} := (n,1,\ldots,1)$, e. g. For the interpolation error with respect to the L_2- or to the L_∞-norm, (19) and (25) lead to

$$\begin{aligned} \gamma_{\mathbf{l},\infty} &= \frac{2^{-4\cdot|\mathbf{l}|_1}\cdot|u|_\infty^2}{4^d\cdot 2^{|\mathbf{l}-\mathbf{1}|_1}} = \frac{1}{2^d}\cdot 2^{-5\cdot|\mathbf{l}|_1}\cdot|u|_\infty^2\,, \\ \gamma_{\mathbf{l},2} &= \frac{2^{-4\cdot|\mathbf{l}|_1}\cdot|u|_2^2}{9^d\cdot 2^{|\mathbf{l}-\mathbf{1}|_1}} = \left(\frac{2}{9}\right)^d\cdot 2^{-5\cdot|\mathbf{l}|_1}\cdot|u|_2^2\,. \end{aligned} \tag{37}$$

Thus, if we use all subspaces $W_{\mathbf{l}}$ for approximation where $\gamma_{\mathbf{l},||.||} \geq \sigma_{||.||}(n)$ with

$$\sigma_\infty(n) \;:=\; \gamma_{\bar{\mathbf{l}},\infty} \;=\; \frac{1}{2^d}\cdot 2^{-5\cdot(n+d-1)}\cdot|u|_\infty^2$$

and

$$\sigma_2(n) \;:=\; \gamma_{\bar{\mathbf{l}},2} \;=\; \left(\frac{2}{9}\right)^d\cdot 2^{-5\cdot(n+d-1)}\cdot|u|_2^2\,,$$

respectively, we obtain the relation $|\mathbf{l}|_1 \leq n+d-1$. This just leads to the definition of the standard sparse grid space $V_n^{(1)}$ in (29). Therefore, our heuristic sparse grid

approach of the previous section is optimal in this cost-benefit setting with respect to both the L_∞- and the L_2-norm.
Now, if we base the approach (36) on the energy norm, (19) and (25) result in

$$\gamma_{\mathbf{l},E} \;=\; \frac{3}{6^d} \cdot 2^{-5\cdot|\mathbf{l}|_1} \cdot \sum_{j=1}^{d} 4^{l_j} \cdot |u|_\infty^2 \tag{38}$$

or a corresponding formula involving $|.|_2$. Since there is nothing different with respect to the order terms (i.e., the terms containing $|\mathbf{l}|_1$), we restrict ourselves to the case of (38). Thus, following the rule $\gamma_{\mathbf{l},E} \geq \sigma_E(n)$ with

$$\sigma_E(n) \;:=\; \gamma_{\bar{\mathbf{l}},E} \;=\; \frac{3}{6^d} \cdot 2^{-5\cdot(n+d-1)} \cdot \left(4^n + 4\cdot(d-1)\right) \cdot |u|_\infty^2 \,,$$

we can define a new sparse grid space $V_n^{(E)}$ which is based on the energy norm:

$$V_n^{(E)} \;:=\; \bigoplus_{|\mathbf{l}|_1 - \frac{1}{5}\cdot\log_2\left(\sum_{j=1}^d 4^{l_j}\right) \;\leq\; (n+d-1)-\frac{1}{5}\cdot\log_2(4^n+4d-4)} W_{\mathbf{l}} \,. \tag{39}$$

Lemma 9 *The sparse grid space $V_n^{(E)}$ is a subspace of $V_n^{(1)}$, and its dimension fulfills*

$$|V_n^{(E)}| \;\leq\; 2^n \cdot \frac{d}{2} \cdot e^d \;=\; O(2^n) \;=\; O(h_n^{-1}) \,. \tag{40}$$

*P*roof: First, for subspaces $W_{\mathbf{l}}$ with $|\mathbf{l}|_1 = n+d-1+i,\ i \in \mathbb{N}$, we have

$$\begin{aligned}
|\mathbf{l}|_1 - \frac{1}{5}\cdot\log_2\left(\sum_{j=1}^d 4^{l_j}\right) \;&\geq\; n+d-1+i-\frac{1}{5}\cdot\log_2\left(4^{n+i}+4d-4\right)\\
&\geq\; n+d-1+i-\frac{1}{5}\cdot\log_2\left(4^i\left(4^n+4d-4\right)\right)\\
&>\; n+d-1-\frac{1}{5}\cdot\log_2\left(4^n+4d-4\right)\,.
\end{aligned}$$

Therefore, no $W_{\mathbf{l}}$ with $|\mathbf{l}|_1 > n+d-1$ can belong to $V_n^{(E)}$. Consequently, we get $V_n^{(E)} \subseteq V_n^{(1)}$ and $|V_n^{(E)}| \;\leq\; |V_n^{(1)}|$ for all $n \in \mathbb{N}$. Starting from that, (19) provides

$$\begin{aligned}
|V_n^{(E)}| \;&=\; \sum_{i=0}^{n-1} \;\sum_{|\mathbf{l}|_1=n+d-1-i,\ \sum_{j=1}^d 4^{l_j}\geq\frac{4^n+4d-4}{32^i}} |W_{\mathbf{l}}|\\
&=\; 2^n\cdot\frac{1}{2}\cdot\sum_{i=0}^{n-1} 2^{-i}\cdot \sum_{|\mathbf{l}|_1=n+d-1-i,\ \sum_{j=1}^d 4^{l_j}\geq\frac{4^n+4d-4}{32^i}} 1\\
&\leq\; 2^n\cdot\frac{1}{2}\cdot\lim_{n\to\infty}\sum_{i=0}^{n-1} 2^{-i}\cdot \sum_{|\mathbf{l}|_1=n+d-1-i,\ \sum_{j=1}^d 4^{l_j}\geq\frac{4^n+4d-4}{32^i}} 1\\
&=\; 2^n\cdot\frac{1}{2}\cdot\lim_{n\to\infty}\sum_{i=0}^{n-1} 2^{-i}\cdot d\cdot\binom{d-1-\lfloor 1.5i\rfloor}{d-1}\,,
\end{aligned}$$

since it can be shown that, for $n \to \infty$, our energy-based sparse grid and the grid resulting from the second condition $|\mathbf{l}|_\infty \geq n - \lfloor 2.5i \rfloor$ for the inner sum instead of

$$\sum_{j=1}^d 4^{l_j} \geq \frac{4^n+4d-4}{32^i}$$

are the same, and since there exist

$$\binom{d-1+\lfloor 1.5i \rfloor}{d-1}$$

such subspaces $W_{\mathbf{l}}$ with $|\mathbf{l}|_\infty = l_1$. Consequently, we obtain

$$\begin{aligned} |V_n^{(E)}| &\leq 2^n \cdot \frac{d}{2} \cdot \sum_{i=0}^{\infty} 2^{-\frac{2}{3}i} \cdot \binom{d-1+i}{d-1} \\ &= 2^n \cdot \frac{d}{2} \cdot \left(1 - 2^{-\frac{2}{3}}\right)^{-d} \\ &\leq 2^n \cdot \frac{d}{2} \cdot e^d , \end{aligned}$$

since $\sum_{i=0}^{\infty} x^i \cdot \binom{k+i}{k} = (1-x)^{-k-1}$ for $k \in \mathbb{N}_0$ and $0 < x < 1$. $\square$

Note that, apart from providing an upper bound on $|V_n^{(E)}|$ for arbitrary n and d, (40) indicates the asymptotic behaviour for $n \to \infty$ and $d \to \infty$. In this asymptotic sense, (40) is sharp.

Table 1 compares the dimensions of the standard full grid approximation space $V_n^{(\infty)}$ with both sparse grid spaces $V_n^{(1)}$ and $V_n^{(E)}$ for $d \in \{2,3,4\}$ and $n \in \{10, 20\}$. The sparse grid effect is obvious already for $V_n^{(1)}$. For higher values of d, the advantages of $V_n^{(E)}$ become evident.

	$d=2$		$d=3$		$d=4$	
	$n=10$	$n=20$	$n=10$	$n=20$	$n=10$	$n=20$
$V_n^{(\infty)}$	$1.05 \cdot 10^6$	$1.10 \cdot 10^{12}$	$1.07 \cdot 10^9$	$1.15 \cdot 10^{18}$	$1.10 \cdot 10^{12}$	$1.21 \cdot 10^{24}$
$V_n^{(1)}$	9217	$1.99 \cdot 10^7$	47103	$2.00 \cdot 10^8$	$1.78 \cdot 10^5$	$1.41 \cdot 10^9$
$V_n^{(E)}$	3841	$4.72 \cdot 10^6$	10495	$1.68 \cdot 10^7$	24321	$5.27 \cdot 10^7$

Table 1: Dimension of $V_n^{(\infty)}$, $V_n^{(1)}$, and $V_n^{(E)}$ for different values of d and n

Theorem 2 *The energy norm of the interpolation error of a function $u \in X_0^p(\Omega)$, $p \in \{2, \infty\}$, in the sparse grid space $V_n^{(E)}$ is bounded by*

$$\begin{aligned} \|u - u_n^{(E)}\|_E &\leq \frac{d \cdot |u|_\infty}{3^{(d-1)/2} \cdot 4^{d-1}} \cdot \left(\frac{1}{2} + \left(\frac{5}{2}\right)^{d-1}\right) \cdot 2^{-n} = O(h_n), \\ \|u - u_n^{(E)}\|_E &\leq \frac{2 \cdot d \cdot |u|_2}{\sqrt{3} \cdot 6^{d-1}} \cdot \left(\frac{1}{2} + \left(\frac{5}{2}\right)^{d-1}\right) \cdot 2^{-n} = O(h_n). \end{aligned} \tag{41}$$

*P*roof: First, since

$$\|u - u_n^{(E)}\|_E \leq \|u - u_n^{(1)}\|_E + \|u_n^{(1)} - u_n^{(E)}\|_E ,$$

and since we know from Theorem 1 that $\|u - u_n^{(1)}\|_E$ is of the order $O(h_n)$, we can restrict ourselves to $\|u_n^{(1)} - u_n^{(E)}\|_E$. For that, it can be shown that, for $i \in \mathbb{N}_0$, each

$W_{\mathbf{l}}$ with $|\mathbf{l}|_1 = n+d-1-i$ and $|\mathbf{l}|_\infty \geq n-2.5i$ is a subspace of $V_n^{(E)}$. Therefore, we obtain with (25)

$$
\begin{aligned}
&||u_n^{(1)} - u_n^{(E)}||_E \\
&\leq \sum_{W_{\mathbf{l}} \subseteq V_n^{(1)} \ominus V_n^{(E)}} ||u_{\mathbf{l}}||_E \\
&\leq \sum_{i=0}^{i^*} \sum_{|\mathbf{l}|_1=n+d-1-i,\ |\mathbf{l}|_\infty<n-2.5i} ||u_{\mathbf{l}}||_E \\
&\leq \frac{|u|_\infty}{2\cdot 12^{(d-1)/2}} \cdot \sum_{i=0}^{i^*} \sum_{|\mathbf{l}|_1=n+d-1-i,\ |\mathbf{l}|_\infty<n-2.5i} 4^{-|\mathbf{l}|_1} \cdot \left(\sum_{j=1}^{d} 4^{l_j}\right)^{1/2} \\
&\leq \frac{|u|_\infty}{2\cdot 12^{(d-1)/2}} \cdot 4^{-n-d+1} \cdot \sum_{i=0}^{i^*} 4^i \cdot \sum_{|\mathbf{l}|_1=n+d-1-i,\ |\mathbf{l}|_\infty<n-2.5i} \left(\sum_{j=1}^{d} 2^{l_j}\right) \\
&\leq \frac{|u|_\infty}{2\cdot 12^{(d-1)/2}} \cdot 4^{-n-d+1} \cdot \sum_{i=0}^{i^*} 4^i \cdot \sum_{j=1}^{n-1-\lfloor 2.5i\rfloor} d \cdot \binom{n+d-2-i-j}{d-2} \cdot 2^j \\
&= \frac{|u|_\infty}{2\cdot 12^{(d-1)/2}} \cdot 4^{-n-d+1} \cdot \sum_{i=0}^{i^*} 4^i \cdot \sum_{k=1}^{n-1-\lfloor 2.5i\rfloor} d \cdot \binom{d-2+\lfloor 1.5i\rfloor+k}{d-2} \cdot 2^{n-\lfloor 2.5i\rfloor-k} \\
&= \frac{d\cdot|u|_\infty}{2\cdot 12^{(d-1)/2}} \cdot 4^{-(d-1)} \cdot 2^{-n} \cdot \sum_{i=0}^{i^*} 2^{-\lfloor \frac{i}{2}\rfloor} \cdot \sum_{k=1}^{n-1-\lfloor 2.5i\rfloor} \binom{d-2+\lfloor 1.5i\rfloor+k}{d-2} \cdot 2^{-k} \\
&\leq \frac{d\cdot|u|_\infty}{2\cdot 12^{(d-1)/2}} \cdot 4^{-(d-1)} \cdot 2^{-n} \cdot 2 \cdot 5^{d-1} \\
&= \frac{d\cdot|u|_\infty}{3^{(d-1)/2}\cdot 4^{d-1}} \cdot \left(\frac{5}{2}\right)^{d-1} \cdot 2^{-n},
\end{aligned}
$$

where $0 \leq i^* \leq n-1$ is the maximum value of i for which the set of indices $\mathbf{l}$ with $|\mathbf{l}|_1 = n+d-1-i$ and $|\mathbf{l}|_\infty < n-2.5i$ is not empty. Together with (34) from Theorem 1, we get the first result and, in a completely analogous way, the second one, too. $\square$

Though we have only derived upper bounds for the energy norm of the interpolation errors, it is helpful to compare the respective results (28) from Lemma 6, (34) from Theorem 1, and (41) for our three discrete approximation spaces $V_n^{(\infty)}$, $V_n^{(1)}$, and $V_n^{(E)}$. Table 2 shows that there is no asymptotic growth with respect to d, neither for the full grid case nor for our two sparse grid spaces.

As the first main result of this section, note that the dimension of the sparse grid space $V_n^{(E)}$ is of the order $O(2^n)$ without any polynomial term in n as we have it for $V_n^{(1)}$ (cf. Lemma 7 and Lemma 9). On the other hand, the interpolation error with respect to the energy norm is of the order $O(2^{-n})$ for both spaces $V_n^{(E)}$ and $V_n^{(1)}$ (cf. Theorem 1 and Theorem 2). Though there is still some d-dependence in the factors that are constant with respect to n, we learn from the following theorem that, at least for the *representation* of functions and at least for the case of $X_0^\infty(\Omega)$ and $|.|_\infty$, we have overcome the curse of dimension.

Theorem 3 *Let $\varepsilon = \varepsilon(h_n)$ denote the accuracy of the interpolant $u_n^{(E)}$ of some $u \in X_0^\infty(\Omega)$ in the energy-based sparse grid space $V_n^{(E)}$. For those ε, computing $u_n^{(E)}$ needs*

$$N(\varepsilon) \ \leq \ C_d \cdot |u|_\infty \cdot \varepsilon^{-1} \tag{42}$$

$V_n^{(\infty)}$	$V_n^{(1)}$	$V_n^{(E)}$
$\frac{d^{3/2}}{2 \cdot 3^{(d-1)/2} \cdot 6^{d-1}}$	$\frac{d}{2 \cdot 3^{(d-1)/2} \cdot 4^{d-1}}$	$\frac{d}{3^{(d-1)/2} \cdot 4^{d-1}} \cdot \left(\frac{1}{2} + \left(\frac{5}{2}\right)^{d-1}\right)$
$\frac{d^{3/2}}{\sqrt{3} \cdot 9^{d-1}}$	$\frac{d}{\sqrt{3} \cdot 6^{d-1}}$	$\frac{2 \cdot d}{\sqrt{3} \cdot 6^{d-1}} \cdot \left(\frac{1}{2} + \left(\frac{5}{2}\right)^{d-1}\right)$

Table 2: d-depending constants in the bounds for $\|u - u_n^{(.)}\|_E$ (multiply with $|u|_\infty \cdot 2^{-n}$ (first row) or $|u|_2 \cdot 2^{-n}$ (second row) to get the respective bounds)

grid points with C_d uniformly bounded in d.

*P*roof: Lemma 9 says

$$N(h_n) \;=\; N(2^{-n}) \;\leq\; 2^n \cdot \frac{d}{2} \cdot e^d \,,$$

and Theorem 2 provides

$$\varepsilon(h_n) \;=\; \varepsilon(2^{-n}) \;\leq\; \frac{d \cdot |u|_\infty}{3^{(d-1)/2} \cdot 4^{d-1}} \cdot \left(\frac{1}{2} + \left(\frac{5}{2}\right)^{d-1}\right) \cdot 2^{-n} \,.$$

Thus, we obtain for $\varepsilon = \varepsilon(h_n)$

$$\begin{aligned} N(\varepsilon) \;&\leq\; \varepsilon^{-1} \cdot \frac{d^2 \cdot |u|_\infty \cdot e^d}{2 \cdot 3^{(d-1)/2} \cdot 4^{d-1}} \cdot \left(\frac{1}{2} + \left(\frac{5}{2}\right)^{d-1}\right) \\ &=\; \varepsilon^{-1} \cdot \frac{d^2 \cdot |u|_\infty \cdot e}{2} \cdot \left(\left(\frac{5 \cdot e}{\sqrt{3} \cdot 8}\right)^{d-1} + \frac{1}{2} \cdot \left(\frac{e}{\sqrt{3} \cdot 4}\right)^{d-1}\right) \\ &=\; \varepsilon^{-1} \cdot |u|_\infty \cdot C_d \end{aligned}$$

with $C_d < 2012$. □

Note that, for $X_0^2(\Omega)$ and $|.|_2$, the corresponding coefficient C_d is about 1.14^d and, thus, still grows exponentially in d.
In the next section, we finally deal with the solution of the Poisson equation.

5 Solution of the Poisson equation

Now, we turn to the Poisson equation $-\Delta u = f$ on the d-dimensional unit cube $\Omega =]0,1[^d$ with homogeneous Dirichlet boundary conditions. The solution u is supposed to be in one of our spaces $X_0^p(\Omega)$, $p \in \{2, \infty\}$. We start from the weak formulation

$$\text{find} \quad u \in V_n \quad \text{such that} \quad a(u,v) = (f,v) \quad \forall\, v \in V_n \tag{43}$$

with an appropriate finite dimensional space V_n (i. e. $V_n^{(\infty)}$, $V_n^{(1)}$, or $V_n^{(E)}$), and we denote the respective solution by u_n^{FE}. The bilinear form $a(\mathrm{cot}, \cdot)$ is given by

$$a(u,v) \;:=\; \int_\Omega \sum_{j=1}^{d} \frac{\partial u}{\partial x_j} \cdot \frac{\partial v}{\partial x_j} \, \mathrm{d}\mathbf{x} \,,$$

and the energy norm (20) is just the norm naturally induced by $a(\cdot,\cdot)$. Then, as usual in the finite element approach [SF], the previous estimates for the interpolation error lead to upper bounds for the error of the finite element solution u_n^{FE}.

Lemma 10 *Let $u_n^{\mathrm{FE},(k)}$ be the finite element solution of (43) in $V_n^{(k)}$, $k \in \{\infty, 1, E\}$. With respect to the energy norm, the error of this approximate solution is bounded by the error of any $v_n^{(k)} \in V_n^{(k)}$, especially by the error of the interpolant $u_n^{(k)}$:*

$$||u - u_n^{\mathrm{FE},(k)}||_E \;\leq\; ||u - u_n^{(k)}||_E \;=\; O(h_n)\,. \tag{44}$$

*P*roof: This follows from the best approximation property of the finite element solution (Céa's lemma, cf. [SF]) and from the bounds for the interpolation error in (28), (34), and (41). □

Lemma 10 provides the desired result for the accuracy of our full and sparse grid finite element solutions $u_n^{\mathrm{FE},(k)} \in V_n^{(k)}$. Now, the remaining issue is the cost of computing $u_n^{\mathrm{FE},(k)}$. Since, for this task, we want to obtain an optimal result with $O(|V_n^{(k)}|)$ operations, we have to concentrate on the solver to be applied.
On uniform grids, i.e. in the space $V_n^{(\infty)}$, multigrid methods [Ha2, Br], multilevel methods [BPX, Os, Gr2], or (pre-)wavelet techniques [DK, GO2, St] are known to be optimal. This means that, in each iteration step of the method, the error is reduced at least by a factor $\zeta < 1$ which is independent of both n and, thus, the problem size $|V_n^{(\infty)}|$. If e_n^{IT} denotes the error after the IT-th iteration step, we get

$$||e_n^{IT}|| \;\leq\; \zeta \cdot ||e_n^{IT-1}|| \;\leq\; \zeta^{IT} \cdot ||e_n^0||\,. \tag{45}$$

Consequently, the number IT of iterations that are necessary to compute an approximation to the discrete solution up to some prescribed relative accuracy θ is $|\log_\zeta \theta|$. Since one iteration step entails $O(|V_n^{(\infty)}|)$ operations, we obtain an overall complexity of $O(|\log_\zeta \theta| \cdot |V_n^{(\infty)}|)$. Of course, for a system of linear equations that stems from a discretization of a PDE, it does not make sense to iterate beyond the discretization error. Therefore, with $\theta = O(h_n)$, an optimal solver results in a complexity of the order $O(n \cdot |V_n^{(\infty)}|)$.
Now, for the sparse grid space $V_n^{(1)}$, the standard multiplicative multigrid approach is less straightforward, and the algorithms must be adjusted properly. For examples of multigrid methods on sparse grids, see [Gr1, GZZ, Pf, He]. Note that there is a close relation to multigrid methods based on multiple coarse grids [Mu], the frequency decomposition multigrid method [Ha3, Ha4], and its derivates [NR] and grandchildren [Den1, Den2, He]. Numerical experiments show that these sparse grid solvers work as expected, i.e., the number of operations per iteration step is proportional to $|V_n^{(1)}|$, and the convergence rate ζ is independent of n. A rigorous convergence theory like for the full grid case should be possible in a straightforward way, but, as far as we know, does not yet exist for these (multiplicative) multigrid methods.
However, things are different for the additive case. In [GO1], the generalization of the additive multilevel BPX-preconditioner [BPX] to $V_n^{(1)}$ is presented, together with a complete convergence analysis. It is shown that the condition number of the suitably preconditioned sparse grid system matrix does not depend on n. Furthermore, the matrix vector product with the system matrix and the action of the preconditioner can be implemented to need $O(|V_n^{(1)}|)$ operations, see [Bu1, GO1, GO2]. Thus, a gradient method or a related solver results in a complexity of $O(|V_n^{(1)}| \cdot n)$, if we iterate up to discretization accuracy $O(h_n)$ with respect to the energy norm.

Note that, for steepest descent, e. g., we get $\zeta = (\Lambda - \lambda)/(\Lambda + \lambda)$ in (45) with Λ, λ being the largest and the smallest eigenvalue of the preconditioned system matrix. For $V_n^{(E)}$, such a straightforward generalization of the BPX-method is not available. In [GO2], we considered a preconditioner based on a prewavelet approach. It consists of a transformation from the prewavelet basis to the hierarchical basis (16), its transpose, and a simple scaling operator. The condition number of the prewavelet transformed system matrix for the full grid space $V_n^{(\infty)}$ has been shown to be bounded by a constant independent of n. Since nothing but a basis transformation is performed, a switch to a subspace like $V_n^{(1)}$ or $V_n^{(E)}$ results just in the restriction to the corresponding minor block of the system matrix with respect to the prewavelet transformed system. Then, from linear algebra, we know that the condition number of any minor block is bounded by a constant independent of n. Finally, note that the action of both the corresponding system matrices and the prewavelet preconditioners can be implemented to need $O(|V_n^{(\infty)}|)$, $O(|V_n^{(1)}|)$, or $O(|V_n^{(E)}|)$ operations, respectively. Thus, for each space $V_n^{(k)}$, $k \in \{\infty, 1, E\}$, we can construct a solver of complexity $O(|V_n^{(k)}| \cdot n)$, if we only iterate until discretization accuracy $O(h_n)$ with respect to the energy norm is reached. Altogether, we can summarize the above discussion in the following

Theorem 4 *The linear system resulting from the discretization of (43) in the space $V_n^{(k)}$, $k \in \{\infty, 1, E\}$, can be solved by an optimal multilevel method (a prewavelet preconditioned gradient method [GO2] with $\zeta < 1$ independent of n, e. g.) up to a relative error of the order $O(h_n)$ (discretization error) with $O(|V_n^{(k)}| \cdot n)$ operations.*

*P*roof: See [GO2] and the arguments above. □

In the final part of this section, we show how one can get rid of the n-term in the complexity estimate of Theorem 4. To this end, we apply a trick which has been used already for some time by numerical analysts and engineers to accelerate the computations: We embed our optimal multilevel solver in a *nested iteration* scheme. The basic idea is to avoid to start the solution process on the finest level n of discretization with an arbitrary start value, but to use a good approximation instead which has been gained before on the next coarser level of discretization by making one or a few iteration steps of the optimal solver there. The recursive extension of this process results in the following algorithm [KD]:

let $\tilde{u}_1^{\mathrm{FE},(k)} = \tilde{u}_1^{0,(k)} \in V_1^{(k)}$ be an (arbitrary) start value on level 1;
for $l := 2$ to n:
  $\tilde{u}_l^{0,(k)} := P_l^{l-1} \tilde{u}_{l-1}^{\mathrm{FE},(k)}$; (start value on level l)
  for $i := 1$ to m_l:
    apply one iteration step of our optimal method on level l;
  the result is $\tilde{u}_l^{\mathrm{FE},(k)}$;
(46)

Here, $P_l^{l-1} : V_{l-1}^{(k)} \to V_l^{(k)}$ denotes a suitable prolongation. Note that this scheme is not an iterative method, but describes a finite process. Since we use an optimal iterative method on each level, the numbers m_l of iteration steps can be chosen to be independent of l, i. e. $m_l := m$ for all levels l.
We now consider a hierarchy of problems (43) on $V_l^{(k)}$, $l = 1, ..., n$, $k \in \{\infty, 1, E\}$. Though, for $V_n^{(1)}$ and $V_n^{(E)}$, this hierarchy is not as straightforward as it is for $V_n^{(\infty)}$, it can be defined via the natural refinement parameter n. On each level l, the exact

solution u is approximated by $u_l^{\mathrm{FE},(k)}$ with a consistency order $O(h_l)$, i.e.

$$||u - u_l^{\mathrm{FE},(k)}||_E \ \leq \ c \cdot h_l \, .$$

Then, by means of the triangle inequality, the solutions on two successive levels of discretization should differ by $O(h_l + h_{l-1})$. This results in the condition

$$||u_l^{\mathrm{FE},(k)} - P_l^{l-1} u_{l-1}^{\mathrm{FE},(k)}||_E \ \leq \ c_1 \cdot h_l \, . \tag{47}$$

Furthermore, if we can bound $||P_l^{l-1}||_E \cdot h_{l-1}/h_l$ from above by a constant independent of l,

$$||P_l^{l-1}||_E \cdot \frac{h_{l-1}}{h_l} \ \leq \ c_2 \, , \tag{48}$$

we can formulate the following lemma (see also [Ha5], Section 10.5.2):

Lemma 11 *Let m be sufficiently large such that*

$$c_2 \cdot \zeta^m \ < \ 1 \tag{49}$$

is fulfilled. Then, the nested iteration (46) produces $\tilde{u}_l^{\mathrm{FE},(k)}$ for all $l, 1 \leq l \leq n$, with error bounds

$$||\tilde{u}_l^{\mathrm{FE},(k)} - u_l^{\mathrm{FE},(k)}|| \ \leq \ c_3(\zeta^m) \cdot c_1 \cdot h_l \tag{50}$$

where

$$c_3(\zeta^m) := \frac{\zeta^m}{1 - c_2 \cdot \zeta^m} \, , \tag{51}$$

if the start value $\tilde{u}_1^{\mathrm{FE},(k)}$ fulfills (50) with $l = 1$.

*P*roof: Following [Ha5], Section 10.5.2, we make a complete induction over l. As a prerequisite, (50) is valid for $l = 1$. Now, assume that (50) holds for all levels up to $l-1$. For the error $||\tilde{u}_l^{0,(k)} - u_l^{\mathrm{FE},(k)}||_E$ of the starting value $\tilde{u}_l^{0,(k)}$ on level l, we get

$$\begin{aligned}
||\tilde{u}_l^{0,(k)} - u_l^{\mathrm{FE},(k)}||_E &= ||P_l^{l-1}\tilde{u}_{l-1}^{\mathrm{FE},(k)} - u_l^{\mathrm{FE},(k)}||_E \\
&\leq ||P_l^{l-1}u_{l-1}^{\mathrm{FE},(k)} - u_l^{\mathrm{FE},(k)}||_E + ||P_l^{l-1}(\tilde{u}_{l-1}^{\mathrm{FE},(k)} - u_{l-1}^{\mathrm{FE},(k)})||_E \\
&\leq c_1 \cdot h_l + ||P_l^{l-1}||_E \cdot ||\tilde{u}_{l-1}^{\mathrm{FE},(k)} - u_{l-1}^{\mathrm{FE},(k)}||_E \qquad \text{by (47)} \\
&\leq c_1 \cdot h_l + ||P_l^{l-1}||_E \cdot c_3(\zeta^m) \cdot c_1 \cdot h_{l-1} \qquad \text{by (50)} \\
&\leq c_1 \cdot h_l \cdot \big(1 + c_2 \cdot c_3(\zeta^m)\big) \qquad \text{by (48)}\,.
\end{aligned}$$

According to (45), m iteration steps reduce this initial error to

$$||\tilde{u}_l^{\mathrm{FE},(k)} - u_l^{\mathrm{FE},(k)}||_E \ \leq \ \zeta^m \cdot ||\tilde{u}_l^{0,(k)} - u_l^{\mathrm{FE},(k)}||_E \ ,$$

and, using (51), we finally obtain

$$\begin{aligned}
||\tilde{u}_l^{\mathrm{FE},(k)} - u_l^{\mathrm{FE},(k)}||_E &\leq \zeta^m \cdot c_1 \cdot h_l \cdot \big(1 + c_2 \cdot c_3(\zeta^m)\big) \\
&= c_1 \cdot h_l \cdot \frac{\zeta^m}{1 - c_2 \cdot \zeta^m} \\
&= c_1 \cdot h_l \cdot c_3(\zeta^m)\,.
\end{aligned}$$

□

We learn from (50) that, up to the factor $c_3(\zeta^m)$, the bound for the iteration error (50) is equivalent to the bound for the discretization error (47). Due to our choice of the standard prolongation with $||P_l^{l-1}||_E = 1$ and $h_l = h_{l-1}/2$, (48) is fulfilled for $c_2 = 2$. Thus, we get $c_3(\zeta^m) = \zeta^m/(1 - 2\cdot\zeta^m)$. Then, for any optimal method with contraction number $\zeta < 1/2$, condition (49) is fulfilled already for $m = 1$, i.e., just one single iteration step is necessary on each level l in the nested iteration algorithm. Moreover, for $\zeta = 1/3$, we even obtain $c_3(\zeta^1) = 1$, and the bounds for iteration error and discretization error are the same.
Regarding the overall work count, we stated already that the cost per iteration is proportional to $|V_l^{(k)}|$, $l = 1, ..., n$. Furthermore, the cost involved in the prolongation is also at most proportional to $|V_l^{(k)}|$, $k \in \{\infty, 1, E\}$. Finally, we get the following theorem:

Theorem 5 *The linear system resulting from the discretization of (43) in the space $V_n^{(k)}$, $k \in \{\infty, 1, E\}$, is solved by the nested iteration process (46) with any optimal inner iteration ($\zeta < 1$ independent of l) up to a relative error of the order $O(h_n)$ of the discretization error using $O(|V_n^{(k)}|)$ operations only.*

*P*roof:

$$\sum_{l=1}^{n} |V_l^{(k)}| \le \sum_{l=1}^{n} \begin{cases} c_\infty \cdot 2^{d\cdot l} \\ c_1 \cdot 2^l \cdot l^{d-1} \\ c_E \cdot 2^l \end{cases} \le \begin{cases} c_\infty \cdot 2 \cdot 2^{d\cdot n} \\ c_1 \cdot 2 \cdot 2^n \cdot n^{d-1} \\ c_E \cdot 2 \cdot 2^n \end{cases} = \begin{cases} O(2^{d\cdot n}) & k = \infty \\ O(2^n \cdot n^{d-1}) & k = 1 \\ O(2^n) & k = E \end{cases}$$

Here, c_∞, c_1, and c_E are constants which depend on d (see the respective results in (27), Lemma 7, and Lemma 9), but not on n. Thus, the overall cost is of the order $O(|V_n^{(k)}|)$, $k \in \{\infty, 1, E\}$. □

Theorem 5 finally shows that, for the solution of the Poisson equation, the ε-complexity of the energy-based sparse grid approach, i.e. the amount of work necessary to reach an accuracy $\varepsilon = O(h_n)$ in $V_n^{(E)}$, is of the order $O(\varepsilon^{-1})$, without any log-terms in ε. However, in contrast to Theorem 3 dealing with the complexity of the problem of approximation, where the constant factor is uniformly bounded in d, there is no result concerning the behaviour of the constant factors in Theorem 5 with respect to d.

6 Discussion of the complexity

Up to now, as already mentioned in the introduction, the smoothness requirements have been supposed to be fulfilled by the *solution* u, i.e. $u \in X_0^p(\Omega)$. Though the spaces themselves are a quite usual choice in this context, especially for $p = 2$, this is a major difference in comparison with related work concerning the ε-complexity of solving the Poisson equation, such as (e.g.) [We]. There, usually, the smoothness of the *data* (i.e. of the right-hand side f) is taken into account. Therefore, the respective results can not be compared directly. However, in order to make things clear, we want to provide at least some links between both points of view.
Concerning the data f, the task is to characterize the space $Y(\Omega)$ in which the right-hand sides f live for all situations of $u \in X^2(\Omega)$. Obviously, we have $Y(\Omega) \supset X^2(\Omega)$. The other way round, concerning the solutions u, we have to define spaces $Z(\Omega)$ that fulfill $u \in Z(\Omega)$ if and only if $f \in X^2(\Omega)$. Since, roughly speaking, u is smoother than f, we get $Z(\Omega) \subset X^2(\Omega)$. Recently, [GK] has provided the result

$$Z(\Omega) \;=\; \left\{u \in L_2(\Omega): \; D^{\boldsymbol{\alpha}+\mathbf{e}_j} u \in L_2(\Omega), \; |\boldsymbol{\alpha}|_\infty \le 2, \; j = 1, \ldots, d\right\}.$$

From

$$Z(\Omega) \subset X^2(\Omega) \subset Y(\Omega),$$

we can learn that smoothness requirements of the type $f \in X^2(\Omega)$ are stronger than supposing $u \in X^2(\Omega)$. Starting from the first ones, [We] obtains an ε-complexity of the order $\Theta(\varepsilon^{-1/2} \cdot |\log \varepsilon|^{\gamma \cdot d})$ for some positive constant γ, whereas the result of this paper is an upper bound on the ε-complexity of the order $O(\varepsilon^{-1})$, which means a worse ε-exponent due to more general regularity assumptions, but there is no d-dependence in the ε-terms. Note that the constant factor still grows in d. For the mere approximation problem in the (more restrictive) space $X^\infty(\Omega)$, however, the constant factor is uniformly bounded in d (cf. Theorem 3).

7 Concluding remarks

In this paper, we showed how the Poisson equation on the d-dimensional unit cube can be solved with a given accuracy ε in a storage and work count complexity of $O(\varepsilon^{-1})$, if the solution u and, thus, the right-hand side f live in appropriate spaces of bounded mixed derivatives. Consequently, except for the constant factor in the case of L_2-bounded derivatives, there is no dependence on the problem's dimension d. The crucial point of our approach is the fact that the log-terms known from the standard sparse grid or Smolyak context can be eliminated by a further sparsening process following an energy-based cost-benefit optimization.
Of course, these results are not restricted to the Poisson case. At least for linear information, the same arguments also hold for more general elliptic PDEs of the type

$$\sum_{j,k=1}^{d} \frac{\partial}{\partial x_k}\Big(a_{k,j}(x) \cdot \frac{\partial}{\partial x_j} u\Big) + \sum_{j=1}^{d} b_j(x) \cdot \frac{\partial}{\partial x_j} u + c(x) \cdot u = f \quad \text{in} \quad]0,1[^d$$

with sufficiently smooth coefficients $a_{k,j}(x)$, $b_j(x)$, and $c(x)$ and with homogeneous Dirichlet boundary conditions due to the fact that, then, the bilinear form a of the weak formulation is H_0^1-elliptic and bounded. For algorithmic details, see (e. g.) [BD]. In this context, note that the energy-based sparse grid space $V_n^{(E)}$ depends on an operator-based norm. Therefore, in contrast to $V_n^{(\infty)}$ or $V_n^{(1)}$, $V_n^{(E)}$ will also depend on the operator. Furthermore, the results from [BD] show how more general domains can be tackled.
Finally, the whole discussion in this paper has been restricted to the case of piecewise linear finite elements. However, following the concept of [BD, Bu3], the sparse grid approach can be generalized to hierarchical polynomial bases of arbitrary polynomial degree p. Then, we need spaces of bounded mixed derivatives up to order $p+1$ in each direction.

References

[Bab] K. I. Babenko, *Approximation by trigonometric polynomials in a certain class of periodic functions of several variables*, Dokl. Akad. Nauk SSSR, 132 (1960), pp. 982–985 (Russian); 672–675 (English translation).

[Bas] G. Baszenski, F.-J. Delvos, S. Jester, *Blending approximation with sine functions*, in Numerical Methods of Approximation 9, D. Braess and L. L. Schumaker, eds., Int. Ser. Num. Math., 105, 1992, pp. 1–19.

[BPX] J. Bramble, J.P. Pasciak, J. Xu, *Parallel multilevel preconditioners*, Math. Comp., 55 (1990), pp. 1–22.

[Br] A. Brandt, *Guide to multigrid development*, in Multigrid Methods, W. Hackbusch and U. Trottenberg, eds., Lecture Notes in Mathematics 960, Springer, Berlin/Heidelberg, 1982.

[Bu1] H.-J. Bungartz, *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung*, Dissertation, Institut für Informatik, TU München, 1992.

[Bu2] H.-J. Bungartz, *An adaptive Poisson solver using hierarchical bases and sparse grids*, in Iterative Methods in Linear Algebra, P. de Groen and R. Beauwens, eds., Elsevier, Amsterdam, 1992, pp. 293–310.

[Bu3] H.-J. Bungartz, *Finite Elements of Higher Order on Sparse Grids*, Habilitationsschrift, Institut für Informatik, TU München, 1998.

[BD] H.-J. Bungartz, T. Dornseifer, *Sparse grids: Recent developments for elliptic partial differential equations*, Rep. TUM-I9702, Institut für Informatik, TU München, 1997; to appear Multigrid Methods V, LNCSE, Springer, 1998.

[BDZ] H.-J. Bungartz, T. Dornseifer, C. Zenger, *Tensor product approximation spaces for the efficient numerical solution of partial differential equations*, to appear in Proc. Int. Workshop on Scientific Computations, Konya, 1996, Nova Science Publishers, 1998.

[DK] W. Dahmen, A. Kunoth, *Multilevel preconditioning*, Numer. Math., 63 (1992), pp. 315–344.

[Del] F.-J. Delvos, *d-Variate Boolean interpolation*, J. Approx. Theory, 34 (1982), pp. 99–114.

[Den1] J. Dendy, *Revenge of the semicoarsening frequency decomposition multigrid method*, in Proc. 7^{th} Copper Mountain Conference on Multigrid Methods, D. Melson, T. A. Manteuffel, S. F. McCormick, and C. C. Douglas, eds., NASA CP 3339, Hampton, VA, 1996, pp. 227–239.

[Den2] J. Dendy, *Grandchild of the frequency decomposition multigrid method*, SIAM J. Sci. Comput., 16 (1995), pp. 307–319.

[Fa] G. Faber, *Über stetige Funktionen*, Math. Ann., 66 (1909), pp. 81–94.

[Go] W. J. Gordon, *Distributive lattices and the approximation of multivariate functions*, in Approximation with Special Emphasis on Spline Functions, I. J. Schoenberg, ed., Academic Press, New York, 1969, pp. 223–277.

[Gr1] M. Griebel, *A parallelizable and vectorizable multi-level algorithm on sparse grids*, in Parallel Algorithms for Partial Differential Equations, Notes on Numerical Fluid Mechanics 31, W. Hackbusch, ed., Vieweg, Braunschweig/Wiesbaden, 1991, pp. 94–100.

[Gr2] M. Griebel, *Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen*, Teubner Skripten zur Numerik, Teubner, Stuttgart, 1994.

[GHZ] M. Griebel, W. Huber, C. Zenger, *Numerical turbulence simulation on a parallel computer using the combination method*, in Flow Simulation on High-Performance Computers II, Notes on Numerical Fluid Mechanics 52, E. H. Hirschel, editor, Vieweg, Braunschweig/Wiesbaden, 1996, pp. 34–47.

[GK] M. Griebel, S. Knapek, *Optimized approximation spaces for operator equations*, Preprint SFB 256, Universität Bonn, 1998.

[GO1] M. Griebel, P. Oswald, *On additive Schwarz preconditioners for sparse grid discretizations*, Numer. Math., 66 (1994), pp. 449–493.

[GO2] M. Griebel, P. Oswald, *Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems*, Advances in Computational Mathematics, 4 (1995), pp. 171-206.

[GSZ] M. Griebel, M. Schneider, C. Zenger, *A combination technique for the solution of sparse grid problems*, in Iterative Methods in Linear Algebra, P. de Groen and R. Beauwens, eds., Elsevier, Amsterdam, 1992, pp. 263–281.

[GZZ] M. Griebel, C. Zenger, S. Zimmer, *Multilevel Gauß-Seidel-algorithms for full and sparse grid problems*, Computing, 50 (1995), pp. 127–148.

[Ha1] W. Hackbusch, *Theorie und Numerik elliptischer Differentialgleichungen*, Teubner Studienbücher Mathematik, Teubner, Stuttgart, 1986.

[Ha2] W. Hackbusch, *Multigrid Methods and Applications*, Springer, Berlin/Heidelberg, 1985.

[Ha3] W. Hackbusch, *The frequency decomposition multi-grid method. I: Application to anisotropic problems*, Numer. Math., 56 (1989), pp. 229–245.

[Ha4] W. Hackbusch, *The frequency decomposition multi-grid method. II: Analysis based on the additive Schwarz method*, Numer. Math., 63 (1992), pp. 433–453.

[Ha5] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Applied Mathematical Sciences, Vol. 95, Springer, New York, 1993.

[He] P. W. Hemker, *Sparse-grid finite-volume multigrid for 3D-problems*, Advances in Computational Mathematics, 4 (1995), pp. 83–110.

[HM] J. P. Hennart, E. H. Mund, *On the h- and p-versions of the extrapolated Gordon's projector with applications to elliptic equations*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 773–791.

[KD] L. Kronsjö, G. Dahlquist, *On the design of nested iterations for elliptic difference equations*, BIT, 11 (1971), pp. 63–71.

[MT] S. Martello, P. Toth, *Knapsack Problems : Algorithms and Computer Implementations*, John Wiley & Sons, Chichester, 1990.

[Mu] W. A. Mulder, *A new multigrid approach to convection problems*, J. Comput. Phys., 83 (1989), pp. 303–323.

[NR] N. Naik, J. van Rosendale, *The improved robustness of multigrid elliptic solvers based on multiple semicoarsened grids*, Report 55-70, ICASE, 1991.

[Os] P. Oswald, *Multilevel Finite Element Approximation: Theory and Applications*, Teubner Skripten zur Numerik, Teubner, Stuttgart, 1994.

[Pf] C. Pflaum, *Diskretisierung elliptischer Differentialgleichungen mit dünnen Gittern*, Dissertation, Institut für Informatik, TU München, 1996.

[Sm] S. A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR. 148 (1963), pp. 1042–1045 (Russian); 240–243 (English translation).

[St] R. Stevenson, *Robustness of the additive and multiplicative frequency decomposition multi-level method*, Report RANA94-22, Eindhoven University of Technology, 1994.

[SF] G. Strang, G. J. Fix, *An Analysis of the Finite Element Method*, Prentice Hall, Englewood Cliffs, 1973.

[Te] V. N. Temlyakov, *On approximate recovery of functions with bounded mixed derivative*, J. Complexity, 9 (1993), pp. 41–59.

[TWW] J. F. Traub, G. W. Wasilkowski, H. Wozniakowski, *Information-Based Complexity*, Academic Press, New York, 1988.

[WW] G. W. Wasilkowski, H. Wozniakowski, *Explicit cost bounds of algorithms for multivariate tensor products*, J. Complexity, 11 (1995), pp. 1–56.

[We] A. C. Werschulz, *The complexity of the Poisson problem for spaces of bounded mixed derivatives*, Technical Report CUCS-016-95, 1995.

[Ys1] H. Yserentant, *On the multi-level splitting of finite element spaces*, Numer. Math., 49 (1986), pp. 379–412.

[Ys2] H. Yserentant, *Hierarchical bases*, in Proc. ICIAM'91, Washington, 1991, R. E. O'Malley, Jr. et al., eds., SIAM, Philadelphia, 1992.

[Ze] C. Zenger, *Sparse grids*, in Parallel Algorithms for Partial Differential Equations, Notes on Numerical Fluid Mechanics 31, W. Hackbusch, ed., Vieweg, Braunschweig, 1991.